



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 123594

TO: Jennifer Graser
Location: REM/3B09/3C18
Art Unit: 1645
June 7, 2004

Case Serial Number: 10/712713

From: P. Sheppard
Location: Remsen Building
Phone: (571) 272-2529

sheppard@uspto.gov

Search Notes

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2004, 04:13:01 ; Search time 135 Seconds
(without alignments)
5417.965 Million cell updates/sec

Title: US-10-712-713-1

Perfect score: 1318

Sequence: 1 atgacacactatcatttgg.....gaaaaatgcgtttaagcctt 1318

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 50

Total number of hits satisfying chosen parameters: 8

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
 - 2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
 - 3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*
 - 4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*
 - 5: /cgn2_6/ptodata/2/ina/PCFUS COMB.seq:*
 - 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1314	99.7	2424	4	US-08-956-171E-392	Sequence 392, App
2	275	20.9	619	4	US-08-936-165A-179	Sequence 179, App
3	90	6.8	888	3	US-08-714-918-61	Sequence 61, Appl
4	90	6.8	888	3	US-09-265-315-61	Sequence 61, Appl
5	90	6.8	888	3	US-09-265-315-61	Sequence 61, Appl
6	90	6.8	888	3	US-09-266-417-61	Sequence 61, Appl
7	90	6.8	888	4	US-09-528-709-61	Sequence 61, Appl
8	90	6.8	888	4	US-09-527-745-61	Sequence 61, Appl

ALIGNMENTS

RESULT 1

US-08-956-171E-392

Sequence 392, Application US/08956171E

Patent No. 6593114

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

Gil H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,171E

FILING DATE: 20-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861

FILING DATE: January 5, 1996

APPLICATION NUMBER: 08/781,986

FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB248P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (240) 314-1224

TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 392:

SEQUENCE CHARACTERISTICS:

LENGTH: 2424 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 392:

US-08-956-171E-392

Query Match 99.7%; Score 1314; DB 4; Length 2424;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACACTATCATCTTCGCGAATTAAGGTTCTGGCATGAGTTCATTAGCACAATC 60

DB 627 ATGACACACTATCATCTTCGCGAATTAAGGTTCTGGCATGAGTTCATTAGCACAATC 686

QY 61 ATGATGATGATTTAGGACATGAAAGTTCAAGGATCGGATATTGAGAACTACGTTATTACAGAA 120

DB 687 ATGATGATGATTTAGGACATGAAAGTTCAAGGATCGGATATTGAGAACTACGTTATTACAGAA 746

QY 121 GTTGCTCTTTAGAAAATAAGGGATAAAAAATATTACCAATTTGATGCTTAATACATAAAGAA 180

DB 747 GTTGCTCTTTAGAAAATAAGGGATAAAAAATATTACCAATTTGATGCTTAATACATAAAGAA 806

QY 181 GATATGATGATTTATACAAAGTAAATGCTTCGCGAGTAGCCATGAAGAAATAGTACGTGCA 240

DB 807 GATATGATGATTTATACAAAGTAAATGCTTCGCGAGTAGCCATGAAGAAATAGTACGTGCA 866

QY 241 CATCAATTCGAAATTAGATGTTGTAAGTATATAATGATTTTATAGGACAGAGTTATTGATCAA 300

DB 867 CATCAATTCGAAATTAGATGTTGTAAGTATATAATGATTTTATAGGACAGAGTTATTGATCAA 926

QY 301 TATACCTTCAGTAGCTGTAACCTGGTGACATGGTAAACCTTCTCAACACAGGTTTATTATCA 360

DB 927 TATACCTTCAGTAGCTGTAACCTGGTGACATGGTAAACCTTCTCAACACAGGTTTATTATCA 986

QY 361 CATGTTATGAAATGCTGATATAAAGACTTCAATTTTAAATTTGATGCTGACAGGTTATGGA 420

DB 987 CATGTTATGAAATGCTGATATAAAGACTTCAATTTTAAATTTGATGCTGACAGGTTATGGA 1046

QY 421 TTGCGCTGAAAGTGAATTTTCGCTTTTGGGCGATGTAATATAGACCTTCACCTTTTAAAGT 480

DB 1047 TTGCGCTGAAAGTGAATTTTCGCTTTTGGGCGATGTAATATAGACCTTCACCTTTTAAAGT 1106

QY 481 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCTCGATTATTTTAA 540

```
Db 1107 TATAAACCTGATTACGCAATATATGACAAATATGATTTCATCATCTGATTAATTTAA 1166
Qy 541 GATATTAAATGATCTTTTGTGATGATTCACAGAAATGACATATGTTTAAAAAGGTATT 600
Db 1167 GATATTAAATGATCTTTTGTGATGATTCACAGAAATGACATATGTTTAAAAAGGTATT 1226
Qy 601 ATTGCTTTGGGTGATGATGAACATCTACGTAAATTTGAAGCAGATGTTTCAATTTATTAT 660
Db 1227 ATTGCTTTGGGTGATGATGAACATCTACGTAAATTTGAAGCAGATGTTTCAATTTATTAT 1286
Qy 661 TATGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTCAGATAAAGGT 720
Db 1287 TATGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTCAGATAAAGGT 1346
Qy 721 ACTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 1347 ACTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1406
Qy 781 GGTGACCATACAGTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 1407 GGTGACCATACAGTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1466
Qy 841 GATGTTTAAATATTAAGAGCATTTAGAAACGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGT 900
Db 1467 GATGTTTAAATATTAAGAGCATTTAGAAACGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1526
Qy 901 GAACTACAAATTCGAATCAAGTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 1527 GAACTACAAATTCGAATCAAGTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1586
Qy 961 AGTGCTACAAATTCGAATCAAGTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 1587 AGTGCTACAAATTCGAATCAAGTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1646
Qy 1021 CAACCAACACATTTCTCTAGAACACAGGCAATTTTAAATGAATTTGCAGAAAGTTTAAGT 1080
Db 1647 CAACCAACACATTTCTCTAGAACACAGGCAATTTTAAATGAATTTGCAGAAAGTTTAAGT 1706
Qy 1081 AAAGCAGATCGTGTATTTTATGTAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1140
Db 1707 AAAGCAGATCGTGTATTTTATGTAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1766
Qy 1141 TTAACCATACAGATTTTATGTAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1200
Db 1767 TTAACCATACAGATTTTATGTAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1826
Qy 1201 ATTAAATGATTTAGAACAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db 1827 ATTAAATGATTTAGAACAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1886
Qy 1261 CAAAAATTAAGAAATGATTTAGATAAATTTAGGCAATGAAAAATGCGTTTAA 1314
Db 1887 CAAAAATTAAGAAATGATTTAGATAAATTTAGGCAATGAAAAATGCGTTTAA 1940
```

RESULT 2

US-08-936-165A-179/c
; Sequence 179, Application US/08936165A
; Patent No. 6348582

GENERAL INFORMATION:

; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; Polypeptides and their Uses

; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,165A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,032
; FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 179:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-936-165A-179

Query Match 20.9%; Score 275; DB 4; Length 619;

Best Local Similarity 99.7%; Pred. No. 1.1e-119;
Matches 325; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 721 ACTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 611 ACTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 552
Qy 781 GGTGACCATACAGTTTAAATGCAATGCTGTAATGCGATTAGTTATTAGAGAAGCTTA 840
Db 551 GGTGACCATACAGTTTAAATGCAATGCTGTAATGCGATTAGTTATTAGAGAAGCTTA 492
Qy 841 GATGTTACAAATATTTAAAGAGCATTAGAAAGCTTTGGTGGTGGTGGTGGTGGTGGTGGT 900
Db 491 GATGTTACAAATATTTAAAGAGCATTAGAAAGCTTTGGTGGTGGTGGTGGTGGTGGTGGT 432
Qy 901 GAACTACAAATGCAAAATCAAGTTTATTGATGATGATGATGATGATGATGATGATGATGAT 960
Db 431 GAACTACAAATGCAAAATCAAGTTTATTGATGATGATGATGATGATGATGATGATGATGAT 372
Qy 961 AGTGCTACAAATGCAAAAGCAGCAAGAAATATTCACATAAAGAGTTGTTGAGGATTTT 1020
Db 371 AGTGCTACAAATGCAAAAGCAGCAAGAAATATTCACATAAAGAGTTGTTGAGGATTTT 312
Qy 1021 CAACCAACACTTTCTCTAGAACACA 1046
Db 311 CAACCAACACTTTCTCTAGAACACA 286
```

RESULT 3

US-08-714-918-61/c
; Sequence 61, Application US/08714918
; Patent No. 6037123
; GENERAL INFORMATION:
; APPLICANT: Benton, Bret
; APPLICANT: Lee, Ving

APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
TITLE OF INVENTION: TARGET GENES
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,918
FILING DATE: September 13, 1996
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,102
FILING DATE: December 22, 1995
APPLICATION NUMBER: 60/003,798
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 222/005

TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 888 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-714-918-61
Query Match 6.8%; Score 90; DB 3; Length 888;
Best Local Similarity 100.0%; Pred. No. 8.6e-33;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1149 ACAAGATTAAATGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCCTATTAAATGT 1208
DB 567 ACAAGATTAAATGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCCTATTAAATGT 508

QY 1209 ATTAGACAATTGATAATGCTGTTATTTT 1238
DB 507 ATTAGACAATTGATAATGCTGTTATTTT 478

RESULT 4
US-09-265-315-61/c
Sequence 61, Application US/09265315
Patent No. 6187541
GENERAL INFORMATION:
APPLICANT: Benton, Bret
APPLICANT: Lee, Ving J.
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
TITLE OF INVENTION: TARGET GENES

NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/265,315
FILING DATE: March 9, 1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/714,918
FILING DATE: September 13, 1996
APPLICATION NUMBER: 60/009,102
FILING DATE: December 22, 1995
APPLICATION NUMBER: 60/003,798
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 240/247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 888 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-09-265-315-61
Query Match 6.8%; Score 90; DB 3; Length 888;
Best Local Similarity 100.0%; Pred. No. 8.6e-33;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1149 ACAAGATTAAATGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCCTATTAAATGT 1208
DB 567 ACAAGATTAAATGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCCTATTAAATGT 508

QY 1209 ATTAGACAATTGATAATGCTGTTATTTT 1238
DB 507 ATTAGACAATTGATAATGCTGTTATTTT 478

RESULT 5
US-09-265-315-61/c
Sequence 61, Application US/09265315
Patent No. 6187541
GENERAL INFORMATION:
APPLICANT: Benton, Bret
APPLICANT: Lee, Ving J.
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
TITLE OF INVENTION: TARGET GENES
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street

STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/265,315
FILING DATE: March 9, 1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/714,918
FILING DATE: September 13, 1996
APPLICATION NUMBER: 60/009,102
FILING DATE: December 22, 1995
APPLICATION NUMBER: 60/003,798
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 240/247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 888 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-265-315-61

Query Match 6.8%; Score 90; DB 3; Length 888;
Best Local Similarity 100.0%; Pred. No. 8.6e-33;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1149 ACAAGATTAAATTGATAAAATTCGAAGTGCATCGTTAAATTAATGAAGATTCCTATTATGT 1208
Db 567 ACAAGATTAAATTGATAAAATTCGAAGTGCATCGTTAAATTAATGAAGATTCCTATTATGT 508

Qy 1209 ATTAGAACAAATTTGATAATGCTGTTATTTT 1238
Db 507 ATTAGAACAAATTTGATAATGCTGTTATTTT 478

RESULT 6
US-09-266-417-61/c
Sequence 61, Application US/09266417
Patent No. 6228588
GENERAL INFORMATION:
APPLICANT: Benton, Bret
APPLICANT: Lee, Ving J.
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
TITLE OF INVENTION: TARGET GENES
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.

ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/266,417
FILING DATE: March 9, 1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/714,918
FILING DATE: September 13, 1996
APPLICATION NUMBER: 60/009,102
FILING DATE: December 22, 1995
APPLICATION NUMBER: 60/003,798
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 240/248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 888 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-266-417-61

Query Match 6.8%; Score 90; DB 3; Length 888;
Best Local Similarity 100.0%; Pred. No. 8.6e-33;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1149 ACAAGATTAAATTGATAAAATTCGAAGTGCATCGTTAAATTAATGAAGATTCCTATTATGT 1208
Db 567 ACAAGATTAAATTGATAAAATTCGAAGTGCATCGTTAAATTAATGAAGATTCCTATTATGT 508

Qy 1209 ATTAGAACAAATTTGATAATGCTGTTATTTT 1238
Db 507 ATTAGAACAAATTTGATAATGCTGTTATTTT 478

RESULT 7
US-09-528-709-61/c
Sequence 61, Application US/09528709
Patent No. 6630303
GENERAL INFORMATION:
APPLICANT: Benton, Bret
APPLICANT: Lee, Ving
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
TITLE OF INVENTION: TARGET GENES
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible

```
;
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/528,709
; FILING DATE: 17-Mar-2000
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,918
; FILING DATE: September 13, 1996
; APPLICATION NUMBER: 60/009,102
; FILING DATE: December 22, 1995
; APPLICATION NUMBER: 60/003,798
; FILING DATE: September 15, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 222/005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 888 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-09-528-709-61

Query Match      6.8%; Score 90; DB 4; Length 888;
Best Local Similarity 100.0%; Pred. No. 8.6e-33;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1149 ACAAGATTAAATGATAAAATGGAAGTGCATCGTTAAATGAAGATTCTATTAAATGT 1208
DB 567 ACAAGATTAAATGATAAAATGGAAGTGCATCGTTAAATGAAGATTCTATTAAATGT 508

QY 1209 ATTAGAACAAATTTGATAATGCTGTTATTTT 1238
DB 507 ATTAGAACAAATTTGATAATGCTGTTATTTT 478

Search completed: June 6, 2004, 07:10:16
Job time : 136 secs

;
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/528,709
; FILING DATE: 17-Mar-2000
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,918
; FILING DATE: September 13, 1996
; APPLICATION NUMBER: 60/009,102
; FILING DATE: December 22, 1995
; APPLICATION NUMBER: 60/003,798
; FILING DATE: September 15, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 222/005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 888 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-09-528-709-61

Query Match      6.8%; Score 90; DB 4; Length 888;
Best Local Similarity 100.0%; Pred. No. 8.6e-33;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1149 ACAAGATTAAATGATAAAATGGAAGTGCATCGTTAAATGAAGATTCTATTAAATGT 1208
DB 567 ACAAGATTAAATGATAAAATGGAAGTGCATCGTTAAATGAAGATTCTATTAAATGT 508

QY 1209 ATTAGAACAAATTTGATAATGCTGTTATTTT 1238
DB 507 ATTAGAACAAATTTGATAATGCTGTTATTTT 478

RESULT 8
US-09-527-745-61/c
; Sequence 61, Application US/09527745
; Patent No. 6638718
; GENERAL INFORMATION:
; APPLICANT: Benton, Bret
; Lee, Ving
; Malouin, Francois
; Martin, Patrick K.
; Schmid, Molly B.
; Sun, Dongxu
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
; TARGET GENES
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/527,745
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2004, 04:34:41 ; Search time 315 Seconds
(without alignments)
5591.218 Million cell updates/sec

Title: US-10-712-713-1
Perfect score: 1318
Sequence: 1 atgacacatcatcttgt.....gaaaaatgcgtttaagctt 1318

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2089662 seqs, 668146292 residues

Word size : 50

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending Patents NA New:
1: /cgn2_6/ptodata/2/pna/PCT NEW COMB.seq:
2: /cgn2_6/ptodata/2/pna/US06 NEW COMB.seq:
3: /cgn2_6/ptodata/2/pna/US07 NEW COMB.seq:
4: /cgn2_6/ptodata/2/pna/US08 NEW COMB.seq:
5: /cgn2_6/ptodata/2/pna/US09 NEW COMB.seq:
6: /cgn2_6/ptodata/2/pna/US10 NEW COMB.seq:
7: /cgn2_6/ptodata/2/pna/US60 NEW COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
------------	-------	-------------	----	-------------

No matches found

Search completed: June 6, 2004, 08:36:08
Job time : 315 secs

Search completed: June 6, 2004, 07:07:51
Job time : 3824 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2004, 04:07:36 ; Search time 3824 Seconds
(without alignments)
10292.457 Million cell updates/sec

Title: US-10-712-713-1
Perfect score: 1318
Sequence: 1 atgacacactatcatttgt.....gaaaaatcggtttaagctt 1318

Scoring table: OLIGO NUC
Gapop_60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 50

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description
No.							

No matches found

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2004, 23:40:14 ; Search time 3824 Seconds
(without alignments)
10292.457 Million cell updates/sec

Title: US-10-712-713-1

Perfect score: 1318

Sequence: 1 atgacacactatcatttgt.....gaaaaatgcgttttaagctt 1318

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmd.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_est1.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	373.6	28.3	2937	28	BH771016 LLMGTg73
C 2	90	6.8	728	28	BH384961 AG-ND-161
C 3	85.8	6.5	584	28	BH400896 AG-ND-145
C 4	81.8	6.2	1101	29	CNS00396 AL063921 Drosophil

5	77.8	5.9	1200	13	BX415878
6	77.8	5.9	1200	13	BX437758
7	75.4	5.7	1101	29	CNS00EVL
8	75	5.7	908	28	AZ548467
9	71.8	5.4	1099	13	BX456575
10	70.4	5.3	949	28	AZ546009
11	69.6	5.3	754	12	BX376728
12	68.6	5.2	1056	13	BX415058
13	68.2	5.2	594	13	BU498120
14	67.8	5.1	1201	9	AL536104
15	67.2	5.1	844	29	BX139987
16	65.8	5.0	905	28	AZ550256
17	65.6	5.0	1200	13	BX415878
18	65.2	4.9	552	9	AU269872
19	64.6	4.9	825	12	BX374097
20	64.6	4.9	1201	13	BX439779
21	64.4	4.9	806	29	CNS04ARE
22	64	4.9	479	28	BH211305
23	63.8	4.8	843	28	AZ551618
24	63.6	4.8	751	12	BX374223
25	63.4	4.8	1200	13	BX437758
26	63.2	4.8	608	12	BX330059
27	62.6	4.7	856	28	BH178274
28	62.6	4.7	856	29	CNS07KML
29	62.2	4.7	987	29	CNS014PQ
30	61.8	4.7	829	29	BX173672
31	61.8	4.7	1061	13	BX437039
32	61.6	4.7	900	28	AZ549980
33	61.4	4.7	1098	13	BX377526
34	61.4	4.7	1201	13	BX366070
35	61.2	4.6	541	29	CNS04KUS
36	61.2	4.6	1391	29	CG754863
37	61	4.6	1201	13	BX406288
38	61	4.6	1201	13	BX461824
39	60.8	4.6	548	28	BH200391
40	60.8	4.6	1201	13	BX355654
41	60.4	4.6	313	28	BH204978
42	60.4	4.6	795	28	AZ528485
43	60.4	4.6	823	28	AZ676218
44	60.4	4.6	843	28	BH139532
45	60.4	4.6	996	29	CNS000FUH

ALIGNMENTS

RESULT 1
BH771016/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BH771016 2937 bp DNA linear GSS 01-MAY-2002
LLMGtag738 MG1363 Random Sequence Tag Library Lactococcus lactis
subsp. cremoris genomic, genomic survey sequence.

BH771016 GI:20373973

GSS.

Lactococcus lactis subsp. cremoris

Lactococcus lactis subsp. cremoris

Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

Lactococcus.

1 (bases 1 to 2937)

Bolotin,A., Ehrlich,S.D. and Sorokin,A.

Studies of genomes of dairy bacteria Lactococcus lactis

Sci. Alimentis (2002) In press

Contact: Sorokin A

Genetique Microbienne

INRA

CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France

Tel: 33 1 34 65 25 16

Fax: 33 1 34 65 25 21

Email: sorokin@jouy.inra.fr

best homologue in strain LL1403 is murC (91%)

Class: shotgun

High quality sequence start: 30

High quality sequence stop: 2907.

FEATURES	source	Location/Qualifiers
	1. .2937	
		/organism="Lactococcus lactis subsp. cremoris"
		/mol_type="genomic DNA"
		/strain="MG1363"
		/sub_species="cremoris"
		/db_xref="taxon:11359"
		/clone_lib="MG1363 Random Sequence Tag Library"
		/notes="Vector: pSGM02; Site 1: SmaI; Library of chromosomal fragments of L.lactis strain MG1363 was prepared by partial AluI digestion or by sonication."
ORIGIN		
Query Match	28.3%;	Score 373.6; DB 28; Length 2937;
Best Local Similarity	59.4%;	Pred. No. 4e-68;
Matches 671;	Conservative 0;	Mismatches 454; Indels 5; Gaps 2;
Qy	10	TATCATTTTCTCGGAATAAAGGTTCTGGCATGATGTTTCATTAGCACAAATCATGTCATGAT 69
Db	1174	TATCATTTTACCAGTATTAAGGTTTACCGAATGAGTGACCTTCGATGATGTTTACATCAA 1115
Qy	70	TTAGGCATCAAGGTTCAAGGATCGGATATTTGAGAACTACGTATTTACAGAAAGTTGCTCTT 129
Db	1114	ATGGCAAAACGTTCAAGGTTTCAGATTCACACAGACTACTTTTTTACACACGTTGGTCTT 1055
Qy	130	AGAAATAAGGGGATAAAATAATTTACCAATTTGATGCTAATAACATAAAGAGAGATATGGTA 189
Db	1054	GAACAAGCGGGTGTTCCACTTTTGCCGTTTGAAGAAAAAATATTAACCTGAGTTTGAA 995
Qy	190	GTTATACAAGGTAAATGCATT---CGCGAGTAGCCATGAAGAAATAGTACGTGCACATCAA 246
Db	994	TTGATTTGTTGGGAATGCTTTTCGTGATGATATTAATGTCGAAATGCTTTCGCGCATAAA 935
Qy	247	TTGAAATTAGATGTTGTAAGTATTAATGATTTTTTTAGACAGATTAATGATCAATATACT 306
Db	934	AATGGCTTCCCAATTCCAAACGTTATCATGAATTTTTTGGGCCATTTTCATGTAAGATTTCCCA 875
Qy	307	TCAGTAGCTGTAAGTGTGCACATGGTAAACCTTCTACAACAGG--TTTATTATCACATG 364
Db	874	AGTATTGGGGTTGCAGCGGCACTCGGTAAGACTTTCACCCAAATGGAATGTTAGGTTTCATG 815
Qy	365	TTATGAATGTGTATAAAAAGACTTCATTTTTAAATTGGTGTATGGCACAGGTATGGGATTCG 424
Db	814	TGAAGTCGAAAATCGGGGACATTTCTTACCTAATTTGGTTATGGAATCGGGCGTGAATG 755
Qy	425	CTGAAAGTGATATATTTCGCTTTTGAGGCAATGTGTAATATAGACGTCACTTTTTAAGTTATA 484
Db	754	TGGAAGTGAATATTTTGTCTTTTGAATCTAATGAGTATGAACGTCAITTTTCATGCCATTC 695
Qy	485	AACCTGATTACGCAATTTATGACAAATATGATTTTCGATCATCTCGATTATTTTAAAGATA 544
Db	694	ATCTTGAATATACCATTATGACAAATATTTGACTTTTGACCCCTGACTATTTTTCGAGGAA 635
Qy	545	TTAATGATGTTTTTTTGATGCATTCCAAGAAATGGCACATAATGTTTAAAAAGGTTATTATG 604
Db	634	TTGAGATGTCACTTCAGCTTTTCAAGCATATGCAATAATTAATTAATAAAGGAATCTTTG 575
Qy	605	CTTGGGTGATGATGAACAATCTACGTAAAAATGAAGCAGATGTTCCCAATTTATTATTATG 664
Db	574	CTTATGTTGAAGATGTGAATTTGGTAAATTTGTCTGTAAAGCGCTATTTTATTATTATG 515
Qy	665	GATTTAAAGATTCGGATGCACATTTATGCTCAAAATATTCAAATTACGGATAAAGTACTG 724
Db	514	GTTTTGAAGCTTAATGATGATTACCGTGCAGAAAACTTGATTAGAGTACACGGGTTTCAT 455
Qy	725	CTTTTGATGTGTATGCGATGGTGAGTTTATTATGATCACTTCCTGTCTCCCAATATGGTG 784
Db	454	CTTTTGATGCCATTTTCGTGTGTGAAAAAATTTGGTCAITTTTGTGTTCTCCTCATGAA 395
Qy	785	ACCATACAGTTTTAAATGCAATTAGCTGTAAATTGGCATTTAGTATTATTAGAAGAGCTAGATG 844
Db	394	AACATAATGTCTGAAATGCTTTATCTGTGTGGCGGTTTGCCATAATCTCTGGTTGGATA 335

QY	845	TTTCAATAATTAAAGACGATTTAGAAACGCTTTGGTGGTCTTAAACCGTTCGTTCAATGAAA	904
Db	334	TGACAGATGTGCTGACACCTCTTCGACTTTTCGAGGAGTAAACCGCTTTTACTCGAGA	275
QY	905	CTACAATTGCAATCAAGTTATTGTAGATGATTATGCACACCATCCAAAGAGAAATTAGTG	964
Db	274	AAAAAGTAGAGAAACAGTAAATTATTGTGAGACTTCGCTCATCATCAACTGAGATTGAAG	215
QY	965	CTACAATTGAAAACAGCACGAAAGAAATATCCACATAAAGAAAGTTGTTCGAGTATTTCAAC	1024
Db	214	CAACATTTGGATCGGCTCGTCAAAATATCTTCACCGTGAATCGTTGCGGTATTCCAAC	155
QY	1025	CACACACTTTCTCTAGACACACAGCGCAATTTTAAATGAATTTGCAGAAAGTTTAAGTTAAAG	1084
Db	154	CTCATACATTTACCAAGAACCAATTCGCTTTTGCTGATGAGTTTGCAGAAAGTTCTTCACCATG	95
QY	1085	CAGATCGTGTATCTTATGTGAAATTTTGGATCAATTAGAGAAATATCT	1134
Db	94	CGGATACTGTTTATCTGCGCAAAATTTATGGCTCAGCTCGTGTAAAGTAGT	45
RESULT 2			
BH384961/c			
LOCUS			
DEFINITION			
AG-ND-161L18.TR ND-TAM Anopheles gambiae genomic clone			
AG-ND-161L18, genomic survey sequence.			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
COMMENT			
FEATURES			
source			
Location/Qualifiers			
1. .728			
/organism="Anopheles gambiae"			
/mol_type="genomic DNA"			
/strain="PEST"			
/db_xref="taxon:7165"			
/clone="AG-ND-161L18"			
/clone_lib="ND-TAM"			
/note="vector: pBECBAC1; Site_1: HindIII"			
Seq primer: M13 Rev			
Class: BAC ends.			
Other GSSs: AG-ND-161L18.TF			
Contact: Brendan J Loftus			
Department of Eukaryotic Genomics			
The Institute for Genomic Research			
9712 Medical Center Dr., Rockville, MD 20850, USA			
Tel: 301 838 0208			
Fax: 301 838 3543			
Email: b.loftus@tigr.org			
This clone is from an A. gambiae BAC library (ND-TAM) provided by			
F.H. Collins and sequenced by The Institute for Genomic Research			
(TIGR). The BAC library was generated from A. gambiae PEST strain			
DNA. All DNA was extracted from newly hatched first instar larvae			
to minimize the inclusion of DNA from microorganisms that inhabit			
the gut. The DNA is derived from mixed sexes of larvae. The BAC			
library was constructed at Texas A&M University BAC Center			
University, College Station, Texas 77843-2123, USA using a HindIII			
partial digest.			
Seq primer: M13 Rev			
Class: BAC ends.			
Location/Qualifiers			
1. .728			
/organism="Anopheles gambiae"			
/mol_type="genomic DNA"			
/strain="PEST"			
/db_xref="taxon:7165"			
/clone="AG-ND-161L18"			
/clone_lib="ND-TAM"			
/note="vector: pBECBAC1; Site_1: HindIII"			
Seq primer: M13 Rev			
Class: BAC ends.			
Location/Qualifiers			
1. .728			
/organism="Anopheles gambiae"			
/mol_type="genomic DNA"			

The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammos in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```
FEATURES
    source
        1..1101
            /organism="Drosophila melanogaster"
            /mol_type="genomic DNA"
            /db_xref="taxon:7227"
            /clone="BACR08K10"
            /clone_lib="RPCI-98"
            /note="end : TET3"

ORIGIN
Query Match      6.2%; Score 81.8; DB 29; Length 1101;
Best Local Similarity 17.7%; Pred. No. 5.7e-07;
Matches 125; Conservative 300; Mismatches 278; Indels 2; Gaps 1;

QY 538 AAGATATTAATGATGTTTTCATGCAATCCAGAAATCGGCATATTAATCTTAAAAAGGT 597
DB 1098 RRWGGDTWDRDTRKDDWTKWTWKDRADRRWAGDADRWADDGAGTWTATTTWWWW 1039

QY 598 ATTATGCTCGGGGATGATGATCAATCAATGATGATGATGATGATGATGATGATGAT 657
DB 1038 WWWATWDTWDDKWWNAATTAATKTTTWTWTRADWAGDRGAGKRDRAATDADGAGR 979

QY 658 TATTATGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTAACGAGTAA 717
DB 978 RDGGRKDKDKDKDDDKGKKKKAKAATKWKDDWDDWDDWDDWDDWDDWDDWDDWDD 919

QY 718 GGTACTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 777
DB 918 DDGAGDKDDGKGDADDTDGTDKDDDDKDKDDWDDKAGTWDGATWAAATDWWWWGWA 859

QY 778 TATGCTGACCATACAGTTTAAATGATGATGATGATGATGATGATGATGATGATGATG 837
DB 858 DADWWTTWDAADDWADDDWADWADWADWADWADWADWADWADWADWADWADWADWADW 799

QY 838 CTAGATGTTACAAATATTAAGAAGCATTTAGAAACGTTTGGTGGTGTGTTAAACGTCG 897
DB 798 RADDKEDAADRDDDAATWTTTTRDTRDDKWKKTDTWTRWAADRTWDRDDDDDRDRAG 739

QY 898 AATGAAACATCAATTCGAAATCAAGTTATGATGATGATGATGATGATGATGATGATG 957
DB 738 TAGRKWRRTWKRWRKRRDTRWDADADDTDARDDRRRGGDDGADGKGGKKTGRKRRRDR 679

QY 958 ATTAGTGCTACAAATGCAACAGCAGCAAGAAATATCCATAAAGAGTGTGTCAGTA 1017
DB 678 TWDRTDWADAAWTTTDTDDDDKDRRRKRRRRRTTARAADWMTWKAWDWAKW 619

QY 1018 TTTCACACACACATCTTCCTAGACACAGGCATTTTAAATGAAATTTGCAGAAAGTTTA 1077
DB 618 DWKTRADRDWRWAADTW--TDARKADRDWAKARAWRDRARDRARDRRTWTKGTTTIA 561

QY 1078 AGTAAAGCAGATCGTGATCTTCTATGTAATTTTGGATCAATTAGAGAAATATCGGC 1137
DB 560 TWTWAAARAANWAWAATTAATWTTTWTWTTTWTWTTTWTWTTTWTWTTTWTWTTTWT 501

QY 1138 GCATTAAACATCAAGATTTAATGATAAAATTTGAAGGTCATCGCTTAATTAATGAAG 1197
DB 500 AAWTAAWAAAAAATAATTTTWTWTTTWTWTTTWTWTTTWTWTTTWTWTTTWTWTTTWT 441

QY 1198 TCTATTAATGATTAGAAACAAATTTGATGATGATGATGATGATGATGATGATGATGAT 1242
DB 441
```

```
DB 440 TTTTWTWTTWAAATATTTTWTWTTTAAAAAANNNNNNNNNNNNNNNNNNNNNNNNN 396

RESULT 5
BX415878      1200 bp      mRNA      linear      EST 15-MAY-2003
LOCUS      BX415878 Homo sapiens THYMUS Homo sapiens cDNA clone CSOCAP008YI04
DEFINITION      5-PRIME, mRNA sequence.
ACCESSION      BX415878
VERSION      BX415878.1 GI:30765550
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 1200)
AUTHORS      Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
COMMENT      Unpublished (2001)
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Paraday Avenue Genoscope sequence ID : CSOCAP008B502QF1.

FEATURES
    source
        1..1200
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CSOCAP008YI04"
            /tissue_type="THYMUS"
            /clone_lib="Homo sapiens THYMUS"
            /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
            with a NotI-oligo(dT) primer. Five prime end enriched,
            double-strand cDNA was digested with Not I and cloned into
            the Not I and EcoRV sites of the pCMVSPORT 6 vector.
            Library was not normalized."

ORIGIN
Query Match      5.9%; Score 77.8; DB 13; Length 1200;
Best Local Similarity 33.9%; Pred. No. 3.9e-06;
Matches 282; Conservative 115; Mismatches 430; Indels 4; Gaps 1;

QY 417 GGGATTCCTCGTAAAGTGAATTTTCGCTTTTGAGGCATGTGAATATAGACGCTCTTTT 476
DB 369 GGGKKGKGGGGGGGWWWWTTTTTTTTTWHMAAADCTTAWTTTWHCCATSTTSCCTTT 428

QY 477 AAGTTATAAACCTGATTACGCAATATGACAAATATTGATTCGATCATCTGATTTATTT 536
DB 429 AMAAATTTTBSSTWATTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTT 488

QY 537 TAAAGATATTAAATGATGTTTGTGATGATTCACAGAAATGGCACAATAATGTTAAAAAGG 596
DB 489 AAATTTTATTTTAAATTTAAATTTWATAWMTTWWAATTTTAAATTTTAAATTTATWATT 548

QY 597 TATTATTCCTGGGCGTGAATGAACATCTACGTAAATATGAAGCAGATGTTCCAAATTA 656
DB 549 TTTTATTAATTTTAAAAAATAATWATTTTWTWTTTWTWTTTWTWTTTWTWTTTWTWTTT 608

QY 657 TTATTATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAAATACCGATAA 716
DB 609 TWTWTTTATTTTAAATTTWTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 668

QY 717 AGGTACTGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 776
DB 669 WATATWTTTATTTAATTTAGATTAATATATATATATTTTATTTATTTATTTATTTATTT 728

QY 777 ATATGCTGACCATACAGTTTTTAAATGCAATGATGATGATGATGATGATGATGATGATG 836
DB 729 WAAAGAAATWTTTAAATACATANATATATATATATATATATATATATATATATATATAT 788
```



```
QY 837 GCTAGATGTTACAAATATTAAAGACGATTAGAAACGTTTGGTGGTGTAAACGTCGTTT 896
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 789 TTTWAAATWACAAATWAAAATTTTAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 848
QY 897 CAATGAACATCAATTTGCAATCAAGTTATTGTAGATGATTATGACACCAATCCAGAGA 956
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 849 AAAAAAATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 908
QY 957 AATTAGTGTACAAATTTGCAATCAAGTTATTGTAGATGATTATGACACCAATCCAGAGA 1016
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 909 TATTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 968
QY 1017 AATTCAACACACACATTTCTCTAGACACACAGGCAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1076
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 969 WWTATATAATATNTATWAAATWAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1028
QY 1077 AAGTAAAGCAGATCGTGTATTCTTATGTAATTTTGGTCAATTTAGAGAAAATCTCGG 1136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1029 AAAAAAATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1088
QY 1137 CGCA----TTAAGATACAGATTTAATTTGATTAATTTGAAGTGTGATCGTTAATTAATG 1192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1089 ATAAATTTTATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1148
QY 1193 AGATTTCTTATTAATTTAGAACAAATTTGATTAATTTGATTAATTTGATTAATTTGATTAATTTGATTA 1243
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1149 TAAWMAATTAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1199
```

RESULT 6

```
BX437758 1200 bp mRNA linear EST 15-MAY-2003
LOCUS BX437758 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008YB01
DEFINITION 5-PRIME, mRNA sequence.
```

BX437758.1 GI:30773605

ACCESSION

BX437758

KEYWORDS

EST.

SOURCE

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1200)

Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Paraday Avenue Genoscope sequence ID : CS0CAP008CA01QP1.

Location/Qualifiers

1. 1200

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0CAP008YB01"

/tissue="THYMUS"

/clone_lib="Homo sapiens THYMUS"

/notes="vector: pCMVSPORT 6; 1st strand cDNA was primed

with a NotI-oligo(dT) primer. Five prime end enriched,

double-strand cDNA was digested with Not I and cloned into

the Not I and EcoRV sites of the pCMVSPORT 6 vector.

Library was not normalized."

Library was not normalized."

Library was not normalized."

Library was not normalized."

Library was not normalized."

Library was not normalized."

Library was not normalized."

Library was not normalized."

Library was not normalized."

Library was not normalized."

```
QY 445 TTTGAGGCGATGGAATATATAGACGTCACCTTTTAAAGTATTAAACCTGATTACGCAATATG 504
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 TTGGGGGDTTWTWTTTAAARRDKTTTWTWAAAAAARATKRGGRGGADAWTAWT 365
QY 505 ACAATATTTGATTTGATCACTCTGATTTTAAAGATATTAAATGATGTTTGTGATGCA 564
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 366 WTWTTTTATWTTTAAADDAKAWADAGAAATTTTWTWPDAGRGARKRGDTTDFRADWWWW 425
QY 565 TTCCAAGAAATGGGCACATAATGTTTAAAAAGGATTTATTGCTTGGGGTGTATGATGAACAT 624
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 426 WTWAAWAWNDKRTTTTWTTKKRAAAAATTTTWTWAAAAAARATTTTWTWAAAAAARATTT 485
QY 625 CTACGTAAAAATTTGAAGCAGATGTTCCAAATTTTATTATGATTTTAAAGATTCGAGTAC 684
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 486 TWRAAAAAAATTTTAAAGRGARGATTTTAAAGRGARGATTTTAAAGRGARGATTTTAAAGRGARG 545
QY 685 ATTTATGCTCAAAATATTCAAAATTTAGGATTAAGGATGCTCTTTTGTATGTTGATGTTGAT 744
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 546 TTTTATTTAAWATTTATATTTTWTWAAAAAASAGAWKATTTTTKTTTATWKRAG 605
QY 745 GGTGAGTTTATGATCACTCTCTGCTCCACCAATATGTTGACCATACAGTTTAAATGCA 804
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 606 GATTTTWTWAAAGDWKWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 656
QY 805 TTAGCTGTAATTTGCGATTTAGTATTATTTAGAGAAAGTAGATGTTTCAAAATTTAAAGAGCA 864
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 657 WWDGARGRTAKATKWTWTWTWAAWAGAAARKERATTTWTTWAAAAAARATTTTAAAGAGCA 716
QY 865 TTAGAAACGTTTGGTGGTGTAAACGCTGTTCAATGAACTCAATTTGCAATTTCAAGTT 924
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 717 GAAGRAAATTTTATKAGARRGATTTTAAWAWTTATRARGARGADTTTTTATATWWTAT 776
QY 925 ATTGTAGATGATTATGACACCATCCAGAGAAATTTAGTGTCTACA-ATTGAAACAGCAGC 983
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 777 TTTTWTWAAAGATDKAAAAAATTTTAAAGAAATTTWAGARAAAKATTTWAAWAA 836
QY 984 AAAGAAATATCCACATAAAGAGTTGTTGCACTATTTCACACACACACACCTTCTCTAGAAC 1043
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 837 AAAAAAARARWAATATATATTTTATATATAKAAAAAATAAATAAATAAATAAATAAATAAATAA 896
QY 1044 ACAGGCAATTTTAAATGAAATTTGCAAGATTTAAGTAAAGCAGATCGTGTATCTTATG 1103
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 897 AAAAAAATAAAAAATTTTATATAAAAAAATTTTAAWAWRAAATAAATAAATAAATAAATAAATAA 956
QY 1104 TGAATTTTGGATCAATTTAGAGAAATTAATCTGGCGCAATTAAACGATCAAGATTTAATGA 1163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 957 AWAATAATTTTAAATATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1014
QY 1164 TAAATTTGAAGTGCATCGTTAAATTAATGAAGATTTCTATTATGATTAGAAACAATTTGA 1223
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1015 ----ATWTATATAGAAATATATATATATATATATATATATATATATATATATATATATATAT 1070
QY 1224 TAATGCTGTTATTTTATTTATGTTGGTGGCAGGTTGATTTTCAAAATTTCAAAATTTCAATTT 1283
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1071 WTTTTTTTTTTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1130
QY 1284 AGATAAATTTAGCATGAAAAA 1304
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1131 AAAAAAATAAAGAAKADAAWAA 1151
```

RESULT 7

CNS000EVL

LOCUS

CNS000EVL

Drosophila melanogaster genome survey sequence T7 end of BAC:

BACR29B23 of RPCL-98 library from Drosophila melanogaster (fruit

fly), genomic survey sequence.

AL069706

AL069706.1 GI:4949849

KEYWORDS

GSS.

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

ORGANISM

GSS.

GSS.

ORIGIN

Query Match

Best Local Similarity

Matches 272; Conservative

5.9%; Score 77.8; DB 13;

31.6%; Pred. No. 3.9e-06;

169; Mismatches 404; Indels

16; Gaps

3;

Neoptera; Endopterygota, Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> the BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR2923"
/clone_lib="RPCI-98"
/note="end : T7"

Query Match 5.7%; Score 75.4; DB 29; Length 1101;
Best Local Similarity 35.8%; Pred. No. 1.3e-05;
Matches 202; Conservative 97; Mismatches 254; Indels 12; Gaps 2;

QY 753 TTATGATCACTTCCTGCTCCATATGCTGACCATACAGTCTTAAATGCATTAGCTGT 812
DB 460 HTWTMMMMWAAATWTAARAAAWATWTAATWAAAWAAWAAWAAWAAWAAWAAW 519
QY 813 AATTCGAGTATTAGTATAGAGAGCTAGATGTTCAAAATATTAAGAGAGCATTAGAAC 872
DB 520 ATTTTWTWATWTAATWTAATWTAATWTAATWTAATWTAATWTAATWTAATWTA 579
QY 873 GTTTCGCTGTTAAACGTCGTTCAATGAACTCAATTCGAAATCAAGTTATCTAGA 932
DB 580 WTATATTATWTAATWTAATWTAATWTAATWTAATWTAATWTAATWTAATWTA 639
QY 933 ---TGATTATGCACACCATCCAGAGAAATTAGTGCTACAATTTGAAACAGCAGAAAGAA 989
DB 640 ATTTAATTATWTTAATAATAATWTTAAATWTTAAATWTTAAATWTTAAATWTTAA 699
QY 990 ATATCCACATAAAGAGTTGGTGCAGTATTTCAACACACATCTTCTCTAGAACACAGGC 1049
DB 700 AAAAAAATAAATAATWTAATWTAATWTAATWTAATWTAATWTAATWTAATWTA 759
QY 1050 ATTTTAAATGATTTGCGAAGTTTAAAGTAAAGCAGATCGTGTATCTTATGTGAAT 1109
DB 760 ATATATATWTTTAAWWTAAWAAWTAATWTAATWTAATWTAATWTAATWTAATW 810
QY 1110 TTTTTCGATCAATTTAGAGAAATCTCGCGCATTAACGATACAGATTTAATTTGATAAAT 1169
DB 811 AWATAWTAWAAWAAWAAWTAATWTAATWTAATWTAATWTAATWTAATWTAATWTA 870
QY 1170 TGAAGGTGTCATCGTTAAATTAATGAAGATCTTAAATGTATTAGAACAAATTTGATATGC 1229
DB 871 AAWAAAAATAWAAWTTTWTWTTTAAWAAWTAATWTAATWTAATWTAATWTAATWTA 930
QY 1230 TGTATTATTTATGCGGTGCGAGTGAATTCAAAAATTAACAAATTCGATATTAGATAA 1289
DB 931 TAAWAAWTTWATTTTATTTAAATWTAATWTAATWTAATWTAATWTAATWTAATWTA 990

QY 1290 ATTAGCATCAAAAATGCGCTTTTAA 1314
DB 991 ATATWTTATWATAATATATTATTAA 1015

RESULT 8
LOCUS AZ548467
DEFINITION ENTBK30TR Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, genomic survey sequence.
ACCESSION AZ548467
VERSION AZ548467.1 GI:11172102
KEYWORDS GSS.
SOURCE Entamoeba histolytica
ORGANISM Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
REFERENCE 1 (bases 1 to 908)
Loftus B., Van Aken S. and Fraser C.
Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library
Unpublished (2000)
JOURNAL Contact: Brendan J Loftus
COMMENT Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Reverse
Class: shotgun
High quality sequence start: 17
High quality sequence stop: 828.

FEATURES
source
1..908
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/notes="Vector: pHS1; Site 1: Bat I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broch cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, I.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barel, Oxford University Press, 1999)."

ORIGIN
Query Match 5.7%; Score 75; DB 28; Length 908;
Best Local Similarity 43.1%; Pred. No. 1.5e-05;
Matches 363; Conservative 0; Mismatches 480; Indels 0; Gaps 0;

QY 426 TGAAGTGATATTTCGCTTTTGAGCATGTGAATATAGACGTCTTTTAAAGTTATAA 485
DB 33 TGAAGATGATGATGATGAAGAAGATGATGATGAAGAAGATGATGATGAAGAAGATG 92
QY 486 ACCTGATTACCAATTATGCAAAATATTGATTCGATCATCTGATTTTAAAGATAT 545
DB 93 TGATGAAGAAGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGATGATGA 152
QY 546 TAATGATGTTTTTCATGCTATTCCTCAAGAAATGGCCATAATGTTAAAAAGGTTATTTCG 605
DB 153 TGAAGACGATGAAGACGACGATGATGAATTAATGAAGATGATGATGAAGAGAGA 212
QY 606 TTGGGGTGATGATGAACATCTACGTAATAATTGAAGACAGATGTTCCAATTTATTATGG 665

JOURNAL
COMMENT

HM1:IMSS sheared DNA library
Unpublished (2000)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjlftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 26
High quality sequence stop: 796.

FEATURES

source
1..849
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOS1; Site 1; Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."

ORIGIN

Query Match 5.3%; Score 70.4; DB 28; Length 849;
Best Local Similarity 46.4%; Pred. No. 0.00014;
Matches 230; Conservative 0; Mismatches 266; Indels 0; Gaps 0;
QY 520 GATCATCTCGATTATTAAAGATATTAATGATGTTTGTGATGATTCGAAGAATAAGCA 579
DB |||||
QY 571 GATGATGATGAATTTGAAATTAGAAGATGAAGAAGATGAAGATGATGATGAAGAAGAT 512
DB |||||
QY 580 CATAAATGTAAGAAGATATTATGCTTGGGGTATGATGAACATCTACGTPAAATGAA 639
DB |||||
QY 511 GATGATGATGAAGATGATGATGAAGATGAAGAAGATGATGATGAAGAAGATGAAGA 452
QY 640 GCAGATGTTCCCAATTTATTTATGATGATTTAAAGATTCGGATGACATTTATGCTCAAAAT 699
DB |||||
QY 451 GAAGATGATGATGATGATGATTTGAATAGAGATGAAGAAGATGAAGAAGATGAAGAAGAT 392
QY 700 ATTCAAAATACGAATAAGGATCTGCTTTTGTGATGATGATGATGATGATGATGATGAT 759
DB |||||
QY 391 GAT 332
QY 760 CACTTCTGCTCCCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 819
DB |||||
QY 331 GAAGATGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 272
QY 820 ATTAGTATTTAGAGAGCTAGATGTTTACAAATATTAAAGATGATGATGATGATGATGAT 879
DB |||||
QY 271 TTGGAATTAGAAGATGAAGAAGATGAAGATGATGATGATGATGATGATGATGATGATGAT 212
QY 880 GGTGTTAAACGTCGTTTCAATGAACCTACAAATTCGAATTCGAATTCGAATTCGAATTC 939
DB |||||
QY 211 GATGAATTTGAATAGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAT 152
QY 940 GCACACCATCCAGAGAAATTTAGTGTACAAATTCGAACAGCAGCAAGAAATATCCACAT 999
DB |||||
QY 151 GAAGAAGAAGAAGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 92

QY 1000 AAAGAAGTTGTTCCAG 1015
DB 91 GATGAAGATGATGAAG 76
RESULT 11
BJ376728/c
LOCUS
DEFINITION BJ376728 Dictyostelium discoideum cDNA library, CF Dictyostelium
discoideum cDNA clone ddc29013 3', mRNA sequence.
ACCESSION BJ376728.1 GI:19286111
VERSION BJ376728.1
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE 1 (bases 1 to 754)
AUTHORS Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
TITLE Full length cDNA of Dictyostelium discoideum at the culmination
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
source
1..754
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddc29013"
/sex="mat A"
/dev_stage="Culmination stage"
/clone_lib="Dictyostelium discoideum cDNA library, CF"
ORIGIN
Query Match 5.3%; Score 69.6; DB 12; Length 754;
Best Local Similarity 46.8%; Pred. No. 0.00021;
Matches 253; Conservative 0; Mismatches 285; Indels 3; Gaps 1;
QY 324 TGCACATCGTAAAACTTCTACAACAGGTTTATTATCATGTTTATGAATGGTGATAAAAA 383
DB |||||
QY 572 TGAAGAAGAAAGAAAGAAATCAAAGAAATCGATGAAGATTTGATAAATGTTAAAGAG 513
QY 384 GACTTCATTTTAAATTTGTTGATGGCAGGATGGGATTCCTGAAAGTATTTTCGC 443
DB |||||
QY 512 ATTATCAATTTAGAAAGAAAGAAATAAATGGGTGATGATTTATGATGAAGAAGAAAT 453
QY 444 TTTTGAGCGATGTAATATAGACGTCACCTTTTAAAGTTATAAACCTGATAGCAATTAT 503
DB |||||
QY 452 TAATAAGAAATTTAATAAGAAATTAGATATTACAGGTAGTGAAGAAATGATGATGA 393
QY 504 GACAAATTCATTTTCGATCATCTCGATTATTTTAAAGATATTAATGATGTTTTCATGC 563
DB |||||
QY 392 TAATTATTAATTAACAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 333
QY 564 ATTCGAAGAATGGCACAATAATGTTTAAAGAGGATTAATTTGCTGGGGTGAATGAACA 623
DB |||||
QY 332 TTATCAAGAAAGAGGATTTGAAGATGAAAGAAACCTGAAATTCAAATTAATAAATAAT 273
QY 624 TCTACGTAATAATGAGCAGATGTTCCAAATTTATTATTATGATTTAAAGATTCGATGA 683
DB |||||
QY 272 TGAAGAGATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 213
QY 684 CATTTATCTCAAAATATTCAAATTTACGATGAAGGATGATGCTCTTTTGTGATGATCT --- 740
DB |||||
QY 212 ATATTATGATAAATGTTAATATAGTTGATATAAATGATCAAAATGTTGAATATATTC 153
DB |||||
QY 741 GGATGGTGAATTTTATGATCACTTCTCTCCCAATATGTTGACCATACACATTTTAAA 800

Dd	656	AAAAAAYATHTATATAWAAATATATWTWAAWAAWAAWAAWAAACATTTTWCAAAATTTT	597
Qy	1048	GCATTTTTAAATGAAATTTCAGAAAGCTTTAAAGTAAGAAGCAGATCGTGTAATTCCTTATGTGA	1107
Dd	596	TTTTTTTTTATATATATTTTTTTTTTTWTATATATMTATAAWATATATATATTTTTTTTTT	537
Qy	1108	ATTTTGGATCAATTAGAGAAAATACTCGGCCAATTAAACGATACAAGATTAAATIGATAAA	1167
Dd	536	TTTTTTTTTTATATATAMMMMEHAAATTTTWTWTWTAAATTTWTATATAAAATTTTTTTTT	477
Qy	1168	ATTGAGAGTGCATCGTTTAAATTAATGAAGATTCATTAATGATTTAGAACAAATTTGATAAT	1227
Dd	476	TWAAANNTTTTTTTTTTAAATATATATWTTTTTTTTATATAWAARAAAAAAWATATATAT	417
Qy	1228	GCTGTTATTTATTTATGGTGCAGGTGATATTCAAAANAATACAAAATGCATAT	1281
Dd	416	MAMWTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATATATATATATAVAAAATATATATATAT	363
RESULT 13	BU498120		
LOCUS	pFESToab93a09.y1 Plasmodium falciparum 3D7 asexual cDNA Plasmodium	594 bp mRNA linear EST 11-SEP-2002	
DEFINITION	falciparum 3D7 cDNA 5' similar to TR:O96253 O96253 HYPOTHETICAL		
ACCESSION	220.3 KD PROTEIN. ;, mRNA sequence.		
VERSION	BU498120.1 GI:22794314		
KEYWORDS	EST.		
SOURCE	Plasmodium falciparum 3D7		
ORGANISM	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
REFERENCE	1 (bases 1 to 594)		
AUTHORS	Tang,K., Cole,R., Chakrabarti,D., Haywood,R., Clifton,S., Pape,D., Marra,M., Hillier,L., Martin,J., Wylie,T., Dante,M., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Jentes,E., Ronko,I., Teagareishvili,R., Belaygorod,L., Franklin,C., Carr,L., Grow,A., Maguire,L., Richer,J., Watkins,J., Kennedy,S., Levinso,D., Waterston,R., Wilson.R. and Sibley,D. WashU Plasmodium EST Project Unpublished (2001) Contact: L. David Sibley WashU Plasmodium EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu Library was constructed by Debopam Chakrabarti DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: L. David Sibley (lsibley@bcmi.wustl.edu), Washington University Seq primer: -40UP from Gibco High quality sequence stop: 422. Location/Qualifiers 1..594 /organism="Plasmodium falciparum 3D7" /mol_type="mRNA" /db_xref="taxon:36329" /lab_host="DH10B (GeneHog, Invitrogen, Inc.)" /clone_lib="Plasmodium falciparum 3D7 asexual cDNA" /note="Vector: pBluescript SK plus; Site_1: EcoRI; Site_2: XhoI; Library was constructed by Debopam Chakrabarti. Total RNA samples were isolated from mixed stage saponin(0.1%)-lysed P. falciparum 3D7 infected erythrocytes by the acidic guanidium-phenol chloroform method. The poly A+ RNA was isolated by the polyAT-Tract mRNA isolation system (Promega, WI) using streptavidin MagneSphere particles. Directional cDNA libraries were constructed by oligo d(T) priming of poly(A)+ RNA (5mg) into EcoRI and XhoI sites of l ZapII vector using the Zap cDNA synthesis kit (Stratagene, CA). The average size of the cDNA inserts in the library was between 1.0 and 1.5kb Clones were mass excised using the Exaseis helper phage		
FEATURES			
source			

(Stratagene), the phagemids were precipitated with PEG 8000 and extracted with phenol/chloroform. Phagemid DNA was electroporated into DH10B cells."

ORIGIN

```
Query Match      5.2%; Score 68.2; DB 13; Length 594;
Best Local Similarity 47.0%; Pred. No. 0.00042;
Matches 211; Conservative 0; Mismatches 238; Indels 0; Gaps 0;

QY 491 ATTACGCAATATGACAAATATGTTTCGATCACTCCCTGATTATTTAAAGATATAAATG 550
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 ATAATAAAATATTTAAAGTAAATGATGTTTAAATAATATAATATAAAGTAATGATA 182
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 551 ATGTTTTTGATGCATTCACAGAAATGGCACATAATGTTTAAAAAGGTAATATTGCTGGG 610
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 ATGATAATATTAAAGTAATGATATGATTAATATTAAAGTAATGATGATAATATTAAAA 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 611 GTGATGATGAACATCTACGTAAATTAAGACAGATGTTCCAAATTTATTTATGATGATTTA 670
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 GTAATGATAATGATAATATTAAAAATAATGATGTTTAAATAATATAATATAATACCTTTTA 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 671 AAGATTCGGATGACATTTATGCTCAAAATATTCAAATACGGATAAAGTACTGCTTTTG 730
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 AAGTGAAACTTATATTACAGTGTAATAATAAATAAACAAGGAAGACAAAGCGAA 362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 731 ATGTGATGTGATGGTGGATTTTATGATCACTCTCTGCTCCACAATATGTTGACCAT 790
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 ATATACAAATGGGAAATACAAATATATATTTATATATATACCAATAAAAAATAATG 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 791 CAGTTTTAAATGCAATAGCTGTAATGCGATGATTTATTTAGAGAGCTAGATGTTACAA 850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 AAAATAATAATGCTTAAATGATGTTATAATAATAATATTAAATGATTTAGATAATCAAG 482
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 851 ATATTAAAGACATTAGAAAGCTTTGGTGGTTTAAAGCTGTTCAATGAACTACAA 910
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 483 AAAAGGTTATACCTAAAGGTGAACATGATAAATGTACAAACAATATAAATGTTTAAATA 542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 911 TTGCAATCAAGTTTATGATGATGAT 939
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 543 ATACATAATATTAATATAATATAT 571
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 14
AL536104      1201 bp      mRNA      linear      EST 31-MAY-2003
LOCUS      CSODF022YC18 5-PRIME, mRNA sequence.
DEFINITION      Homo sapiens FETAL BRAIN Homo sapiens CDNA clone
ACCESSION      AL536104
VERSION      AL536104.2 GI:31260974
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li W.B., Gruber C., Jessee J. and Polayes D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 13, 2001 this sequence version replaced gi:12799597.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact: Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/InvitrogenCorporation1600
Faraday Avenue Genoscope sequence ID : CSODF022BB09QP1.
```

FEATURES

```
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODF022YC18"
```

/tissue_type="FETAL BRAIN"

/dev_stage="Fetal"

/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

```
Query Match      5.1%; Score 67.8; DB 9; Length 1201;
Best Local Similarity 32.1%; Pred. No. 0.00049;
Matches 195; Conservative 112; Mismatches 298; Indels 3; Gaps 1;

QY 458 AATATAGACGCTCACTTTTTAAAGTTATAAACCTGATTACGCAATTTATGCAAAATATTGATT 517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 594 VAMDVNVVNBSSKKKGGKGGKAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 653
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 518 TCGATCATCTGATTTATTTAAAGATTAATGATGTTTGTGATGCAATTCACAGAAATGG 577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 654 ATATTDMWTTATATAADTDATATAAATTTTWTATTTTATTTTWTATTTTATTTTWTATTA 713
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 578 CACATAATGTTTAAAAAAGGTATTATTCTTGGGGTGATGATGAACATCTACGTAATAATG 637
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 714 WAAATATATATAATATTTTWTATTAATAATTTTWTATTAATTTTAAATTTTATA 773
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 638 AAGCAGATGTTTCCAAATTTATTTATGGAATTTAAGATTCGGATGCAATTTATGCTCAA 697
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 774 WTTTATATATATATTTTWTATTTTWTATTAATAATTTTAAATTTTAAATTTTAA 833
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 698 ATATTCAAATTCAGATAAAGGTACTGCTTTGATGATGATGCGATGCGATGCGATTTATG 757
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 834 WTAAAAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 893
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 758 ATCACTTCTGCTCCACAATATGGTGACCATACAGTTTAAATGCAATGCTGTA---A 814
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 894 TTTTATTTTAAATTTTWTATTAATAATTTTAAATTTTAAATTTTAAATTTTAA 953
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 815 TTGCGATTAGTTATTTAGAGAAGCTAGATGTTACAAATATTAAAGAGCAATTTAGAAAGCT 874
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 954 WTANGHTAATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1013
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 875 TTGCTGGTGTATAAAGCGCTGTTCAATGCAACTACAAATTTGCAAAATCAAGTTATTGATG 934
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1014 WTTTWTTTTWTATWTATTAATAAATTTTATTTTATTTTATTTTATTTTATTTTAA 1073
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 935 ATTATGCACACCATCCAGAGAAATTTAGTCAATTTGAAACAGCAGCAAGAAATATC 994
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1074 AWWRTTWTTTTAAAAAATTTTATTTTAAATTTTAAATTTTAAATTTTAA 1133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 995 CACATAAAGAAAGTTGTTGCGAGTATTTCACACACACTTCTCTAGACACAGGCAATTT 1054
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1134 ATATTATTAATTTTATTAATAATTTTATTTTATTTTATTTTATTTATATAATAA 1193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1055 TAAATGAA 1062
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1194 TAAADRW 1201
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 15

```
BX139987
LOCUS      BX139987      844 bp      DNA      linear      GSS 13-MAR-2003
DEFINITION      Danio rerio genomic clone DKEY-99E7, genomic survey sequence.
ACCESSION      BX139987
VERSION      BX139987.1 GI:27971314
KEYWORDS      GSS.
SOURCE      Danio rerio (zebrafish)
ORGANISM      Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 844)
Humphray,S.J., Huckle,E. and Durham,J.L.
```

TITLE Direct Submission
JOURNAL Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Unpublished

COMMENT This sequence was generated from the T7 end of BAC 99E7. 99E7 is part of the Daniokey BAC Library created by R. Plasterk and N.V. Keygene. Further details: http://www.sanger.ac.uk/projects/D_rerio/.

FEATURES
source Location/Qualifiers
1..844
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-99E7"
/tissue_type="Testis"
/note="vector pindigobAC-536"

ORIGIN

Query Match	5.1%	Score 67.2;	DB 29;	Length 844;
Best Local Similarity	50.0%;	Pred. No. 0.00067;		
Matches 168;	Conservative 0;	Mismatches 168;	Indels 0;	Gaps 0;
QY 426	TGAAAGTGATTTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTAAAGTTATAA	485		
Db 367	TAATAATGATAATAATAATAATGATGATAATAATAATAATGATGATAATAATAA	426		
QY 486	ACCTGATTACGCAATTATGACAAATATTGATTCGATCATCCTGATTTTAAAGATAT	545		
Db 427	TAATAATAATGATAATAATAATAATGATGATAATAATAATAATGATAATAATAA	486		
QY 546	TATGATGTTTTTGTGATGCTCCAGAAATGGCACATAATGTTAAAGGTATTATTGC	605		
Db 487	TAATAATAATGATAATAATAATAATGATGATAATAATAATAATGATAATAATAA	546		
QY 606	TTGGGGTGATGCAACATCTACGTAATAATGAAGCAGATGTTCCAAATTATTATTATGG	665		
Db 547	TAATAATAATGATAATAATAATAATGATAATAATAATAATGATAATAATAATAA	606		
QY 666	ATTAAAGATTCGGATGACATTTATGCTCAAAATTTCAAAATTTACGGATAAGGTACTGC	725		
Db 607	TGATAATAATGATAATAATAACATAATAATAATAATGATAATAATAATAATAATGATGA	666		
QY 726	TTTTGATGTTATGCTGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	761		
Db 667	TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	702		

Search completed: June 6, 2004, 02:44:40
Job time : 3829 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 14:32:30 ; Search time 59 Seconds
(without alignments)
2092.767 Million cell updates/sec

Title: US-10-712-713-2

Perfect score: 2275
Sequence: 1 MTHYFVGKSGMSLAQI.....GDIQLQAYLDKLGKNAF 437

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2275	100.0	437	4 AAU00828	Aau00828 S. aureus
2	2275	100.0	444	6 ABM71210	Abm71210 Staphyloc
3	2275	100.0	444	7 ADD15465	Add15465 Staphyloc
4	2275	100.0	444	7 ADD15463	Add15463 Staphyloc
5	2275	100.0	437	6 ABU15856	Abu15856 Protein e
6	2267	99.6	437	4 AAU33801	Aau33801 Staphyloc
7	2267	99.6	444	4 AAU36700	Aau36700 Staphyloc
8	2238	98.4	437	2 AAU87771	Aaw87771 UDP-N-ace
9	1986	87.3	442	5 ABP40288	Abp40288 Staphyloc
10	1985	87.3	437	6 ABU43254	Abu43254 Protein e
11	1470	64.6	436	6 ABU18407	Abu18407 Protein e
12	1438.5	63.2	445	6 ABU29382	Abu29382 protein e
13	1433.5	63.0	456	4 AAU35277	Aau35277 Enterococ
14	1396	61.4	444	6 ABU23727	Abu23727 Protein e
15	1342	59.0	447	5 ABM48762	Abm48762 Listeria
16	1342	59.0	447	6 ABU33064	Abu33064 Protein e
17	1298.5	57.1	443	6 ABM11616	Abm11616 Alloiooc
18	1180	51.9	457	5 ABM55394	Abm55394 Lactococ
19	1163.5	51.1	443	5 ABP27818	Abp27818 Streptoco
20	1139	50.1	442	5 ABP27819	Abp27819 Streptoco
21	1139	50.1	442	6 ABU46471	Abu46471 Protein e
22	1132.5	49.8	443	6 ABU44255	Abu44255 Protein e
23	1119	49.2	444	4 AAU37914	Aau37914 Streptoco
24	1119	49.2	444	6 ABU46149	Abu46149 Protein e
25	1117	49.1	444	6 ABU01979	Abu01979 S. pneumo

26	1067	46.9	215	2 AAW89199	Aaw89199 Partial s
27	1064	46.8	246	6 ABU43391	Abu43391 Protein e
28	1043	45.8	422	2 AAW55120	Aaw55120 Streptoco
29	1043	45.8	422	5 ABP54614	Abp54614 S. pneumo
30	1043	45.8	422	7 ADC45197	Adc45197 S. pneumo
31	967	42.5	291	7 ADC97287	Adc97287 E. faeciu
32	690.5	30.4	461	6 ABU24704	Abu24704 Protein e
33	641.5	28.2	458	6 ABU23758	Abu23758 Protein e
34	585.5	25.7	450	6 ABU25700	Abu25700 Protein e
35	574	25.2	482	6 ABU16741	Abu16741 Protein e
36	574	25.2	492	6 ADA33736	Ada33736 Acinetoba
37	559	24.6	469	6 ABU33503	Abu33503 Protein e
38	548	24.1	506	3 AAY74419	Aay74419 Neisseria
39	548	24.1	506	6 ABU38136	Abu38136 Protein e
40	545	24.0	469	6 ABU37200	Abu37200 Protein e
41	545	24.0	506	3 AAY74417	Aay74417 Neisseria
42	545	24.0	506	6 ABP80545	Abp80545 N. gonorr
43	538	23.6	468	6 ABU19357	Abu19357 Protein e
44	535	23.5	506	3 AAY74418	Aay74418 Neisseria
45	532	23.4	486	6 ABU35149	Abu35149 Protein e

ALIGNMENTS

RESULT 1
AAU00828 standard; protein; 437 AA.

XX AAU00828;

DT 04-JUL-2001 (first entry)

DE S. aureus UDP-N-acetylmuramate:L-alanine ligase, MurC.

XX UDP-N-acetylmuramate:L-alanine ligase; MurC; immunogen; vaccine;
antibody; wound infection; cellulitis; burn infection; eyelid infection;
KW food poisoning; joint infection; neonatal conjunctivitis; osteomyelitis;
KW skin infection; scalded skin syndrome; toxic epidermal necrosis;
KW Ritter's disease; Lyell's disease; toxic shock syndrome; endocarditis.

XX Staphylococcus aureus.

XX Key Location/Qualifiers
FT Region 137..139

FT /label= Antigenic_epitope

FT Region 236..239

FT /label= Antigenic_epitope

XX WO200116292-A2.

XX 08-MAR-2001.

XX 31-AUG-2000; 2000WO-US023773.

XX 01-SEP-1999; 99US-0151933P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Choi GH;

XX WPI: 2001-183259/18.

XX N-PSDB; AAS00189.

XX New isolated nucleic acid for use in diagnosing Staphylococcus infections
and in vaccines for eliciting immune responses to the infections.

XX Claim 9; Page 14; 225pp; English.

XX The sequence represents S. aureus MurC (UDP-N-acetylmuramate:L-alanine
ligase). The polynucleotides of the invention are used to detect
CC Staphylococcus nucleic acids in a biological sample from an animal for
diagnosing Staphylococcus infections. The polypeptides of the invention

are used to detect anti-Staphylococcus antibodies in a biological sample from an animal to diagnose Staphylococcus infections. The polypeptides are also used in vaccines to elicit protective antibodies in an animal to a member of the Staphylococcus genus and for preventing or attenuating an infection caused by a member of the Staphylococcus genus e.g. wound infection, cellulitis, burn infection, eyelid infection, food poisoning, joint infection, neonatal conjunctivitis, osteomyelitis, skin infection, scalded skin syndrome (also known as toxic epidermal necrosis, Ritter's disease and Lyell's disease), toxic shock syndrome and endocarditis. The polynucleotides may also be used in vaccines and for preventing or attenuating a Staphylococcus infection. Antibodies to the polypeptides may be used to purify, detect and target the polypeptides in vitro and in vivo diagnostic and therapeutic methods

XX SQ Sequence 437 AA;

Query Match 100.0%; Score 2275; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 9.7e-193;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTHYFVGIKSGMSLAQIMHDLGHEVQSGSDIENVVTEVALRNKGIKILPFDANNIKE 60
DB 1 MTHYFVGIKSGMSLAQIMHDLGHEVQSGSDIENVVTEVALRNKGIKILPFDANNIKE 60

QY 61 DMVVIQGNFASHEEIVRAHQKLDVSVYNDFLGQIIDQYTSVAVTGAGKSTTTGLLS 120
DB 61 DMVVIQGNFASHEEIVRAHQKLDVSVYNDFLGQIIDQYTSVAVTGAGKSTTTGLLS 120

QY 121 HVMNGDKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDPDYFK 180
DB 121 HVMNGDKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDPDYFK 180

QY 181 DINDVDFQFQEMAHNVKGGIIAWGDDHLRKTADVPYIYYGFKSDDDIYAQNIQITDKG 240
DB 181 DINDVDFQFQEMAHNVKGGIIAWGDDHLRKTADVPYIYYGFKSDDDIYAQNIQITDKG 240

QY 241 TAFDVYDGEFFYDHFSLSPQYGDHTVINALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
DB 241 TAFDVYDGEFFYDHFSLSPQYGDHTVINALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300

QY 301 ETTIANQVIVDDYAHHPREISATITETARKKYPHKEVVAVFQPHFTFSRTQAFINEFAESLS 360
DB 301 ETTIANQVIVDDYAHHPREISATITETARKKYPHKEVVAVFQPHFTFSRTQAFINEFAESLS 360

QY 361 KADRVFLCEIFGSIKRENTGALTQDLIDKIEGASLINESINLVLEQFONAVILFMGAGDI 420
DB 361 KADRVFLCEIFGSIKRENTGALTQDLIDKIEGASLINESINLVLEQFONAVILFMGAGDI 420

QY 421 OKLQNAVLDKLGKNAF 437
DB 421 OKLQNAVLDKLGKNAF 437

RESULT 2
ABM71210
ID ABM71210 standard; protein; 444 AA.

XX AC ABM71210;

XX DT 20-NOV-2003 (first entry)

XX DE Staphylococcus aureus protein #450.

XX KW Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
XX KW enzymatic assay; antibiotic target.

XX OS Staphylococcus aureus.

XX PN WO200294868-A2.

XX PD 28-NOV-2002.

XX PF 27-MAR-2002; 2002WO-1B002637.

XX 27-MAR-2001; 2001GB-00007661.
XX (CHIR-) CHIRON SPA.
XX Masignani V, Mora M, Scarselli M;
XX WPI; 2003-120786/11.
XX N-PSDB; ACF72770.
XX New Staphylococcus aureus protein, useful as a vaccine for treating or preventing Staphylococcal infection, specifically an infection caused by S. aureus, e.g. sepsis.
XX Claim 1; SEQ ID NO 900; 49pp; English.
XX The invention relates to novel genes and encoded proteins from Staphylococcus aureus. A composition comprising the S. aureus protein, a nucleic acid encoding the protein, or an antibody to the protein, is useful as a pharmaceutical, particularly as a vaccine for treating or preventing infection due to Staphylococcus bacteria, specifically an infection caused by S. aureus. The composition is particularly useful for treating or preventing sepsis in a patient. The composition can also be used for diagnostics. The protein is also used in an assay for enzymatic studies and as a target for antibiotics. This sequence represents one of the novel S. aureus proteins of the invention

XX SQ Sequence 444 AA;

Query Match 100.0%; Score 2275; DB 6; Length 444;
Best Local Similarity 100.0%; Pred. No. 9.9e-193;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTHYFVGIKSGMSLAQIMHDLGHEVQSGSDIENVVTEVALRNKGIKILPFDANNIKE 60
DB 8 MTHYFVGIKSGMSLAQIMHDLGHEVQSGSDIENVVTEVALRNKGIKILPFDANNIKE 67

QY 61 DMVVIQGNFASHEEIVRAHQKLDVSVYNDFLGQIIDQYTSVAVTGAGKSTTTGLLS 120
DB 68 DMVVIQGNFASHEEIVRAHQKLDVSVYNDFLGQIIDQYTSVAVTGAGKSTTTGLLS 127

QY 121 HVMNGDKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDPDYFK 180
DB 128 HVMNGDKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDPDYFK 187

QY 181 DINDVDFQFQEMAHNVKGGIIAWGDDHLRKTADVPYIYYGFKSDDDIYAQNIQITDKG 240
DB 188 DINDVDFQFQEMAHNVKGGIIAWGDDHLRKTADVPYIYYGFKSDDDIYAQNIQITDKG 247

QY 241 TAFDVYDGEFFYDHFSLSPQYGDHTVINALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
DB 248 TAFDVYDGEFFYDHFSLSPQYGDHTVINALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 307

QY 301 ETTIANQVIVDDYAHHPREISATITETARKKYPHKEVVAVFQPHFTFSRTQAFINEFAESLS 360
DB 308 ETTIANQVIVDDYAHHPREISATITETARKKYPHKEVVAVFQPHFTFSRTQAFINEFAESLS 367

QY 361 KADRVFLCEIFGSIKRENTGALTQDLIDKIEGASLINESINLVLEQFONAVILFMGAGDI 420
DB 368 KADRVFLCEIFGSIKRENTGALTQDLIDKIEGASLINESINLVLEQFONAVILFMGAGDI 427

QY 421 OKLQNAVLDKLGKNAF 437
DB 428 OKLQNAVLDKLGKNAF 444

RESULT 3

ADD15465

XX ID ADD15465 standard; protein; 444 AA.

XX AC ADD15465;

XX DT 15-JAN-2004 (first entry)

XX Staphylococcus aureus MurC protein (SeqID 30).
DE microbial; antimicrobial; membrane biosynthesis; pathogenic;
XX immunological response; vaccination; surface disinfectant;
KW personal hygiene application; food preservative; enzyme; MurC;
KW UDP-N-acetylmuramate-alanine ligase.
XX
XX Staphylococcus aureus.
OS
XX WO2003025007-A2.
PN
XX 27-MAR-2003.
PD
XX
XX 20-SEP-2002; 2002WO-CA001428.
PF
XX 21-SEP-2001; 2001US-0323992P.
XX 21-SEP-2001; 2001US-0324152P.
PR 25-SEP-2001; 2001US-0324692P.
PR 26-OCT-2001; 2001US-0339924P.
PR 29-OCT-2001; 2001US-0350973P.
PR 30-OCT-2001; 2001US-0340924P.
PR 27-NOV-2001; 2001US-0333666P.
PR 18-DEC-2001; 2001US-0341732P.
PR 18-DEC-2001; 2001US-0341776P.
PR 19-DEC-2001; 2001US-0341949P.
XX
XX (AFFI-) AFFINIUM PHARM INC.
PA
XX Edwards A, Dharamsi A, Vedadi M, Alam MZ, Awrey D, Beattie B;
PI Domagala M, Houston S, Kanagarajah D, Nethery K, Ng I, Mansoury K;
PI McDonald M, Pinder B, Viola C, Wrezel O;
XX
XX WPI; 2003-468119/44.
DR N-PSDB; ADD15464.
XX
XX Novel crystallized recombinant polypeptides from Staphylococcus aureus,
PT Streptococcus pneumoniae and Escherichia coli and which are involved in
membrane biosynthesis, useful as targets for pathogenic bacteria.
XX
XX Claim 20; SEQ ID NO 30; 325pp; English.
PS
XX This invention relates to the structural and functional characterisation
CC of microbial polypeptides from Staphylococcus aureus (S. aureus),
CC Streptococcus pneumoniae (S. pneumoniae) and Escherichia coli (E. coli)
CC that provide novel antimicrobial targets. Specifically, it refers to
CC polypeptides that are involved in membrane biosynthesis, which play a
CC critical role in the life cycle and viability of their pathogenic species
CC of origin, and hence provide valuable drug targets. Furthermore, the
CC invention describes modified version of these proteins that facilitate
CC characterisation by labelling with isotopic or heavy atoms, and also
CC fusion proteins. These proteins provide structural and functional
CC information to aid the discovery of therapeutic molecules to treat
CC disorders associated with a particular pathogenic species. As such, they
CC are useful for inducing an immunological response in an individual and as
CC an antigen for vaccination purposes. The polypeptides are also useful for
CC developing antimicrobial agents for use as surface disinfectants,
CC personal hygiene applications and as food preservatives or in treating
CC food products to eliminate potential pathogens. This polypeptide sequence
CC is the MurC protein encoded by experimentally predicted DNA from S.
CC aureus UDP-N-acetylmuramate-alanine ligase of the invention.
XX
XX Sequence 444 AA;
XX
XX Query Match 100.0%; Score 2275; DB 7; Length 444;
XX Best Local Similarity 100.0%; Pred. No. 9.9e-193;
XX Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX

QY 1 MTHYFVGIKSGMSSLAQIMHDLGHEVQGSIDIENYVTEVALRNKGKILPFDANNIKE 60
DB 8 MTHYFVGIKSGMSSLAQIMHDLGHEVQGSIDIENYVTEVALRNKGKILPFDANNIKE 67
QY 61 DMVVIQGNFASHEEIVRAHQKLDVVSYNDFLGQIIDQYTSVAVTGANGKSTTGLLS 120

DB 68 DMVVIQGNFASHEEIVRAHQKLDVVSYNDFLGQIIDQYTSVAVTGANGKSTTGLLS 127
QY 121 HVMNGDKKTSFLIGDGTGMLPESDYFAFEACEYRRHFLSYKPDYAIMTIDFDPDYFK 180
DB 128 HVMNGDKKTSFLIGDGTGMLPESDYFAFEACEYRRHFLSYKPDYAIMTIDFDPDYFK 187
QY 181 DINDVDFAFQEMAHNVKGGIIAWGDDEHURKTEADVPYIYGGFKSDSDDIYAQNIQITDKG 240
DB 188 DINDVDFAFQEMAHNVKGGIIAWGDDEHURKTEADVPYIYGGFKSDSDDIYAQNIQITDKG 247
QY 241 TAFDYYVDGEFYDHFSLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKREN 300
DB 248 TAFDYYVDGEFYDHFSLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKREN 307
QY 301 ETTIANQVIVDDYAHHPREISATITARKKYPHKEVAVFQHPFTSKTOAFINEFAESLS 360
DB 308 ETTIANQVIVDDYAHHPREISATITARKKYPHKEVAVFQHPFTSKTOAFINEFAESLS 367
QY 361 KADRVFLCEIFGSIKRENTGALTQDLIDKIEGASLINESINLVQFDNAVILFMGAGDI 420
DB 368 KADRVFLCEIFGSIKRENTGALTQDLIDKIEGASLINESINLVQFDNAVILFMGAGDI 427
QY 421 QKLNAYLDKLGKNAF 437
DB 428 QKLNAYLDKLGKNAF 444
XX
XX RESULT 4
XX ADD15463
XX ID ADD15463 standard; protein; 444 AA.
XX AC ADD15463;
XX DT 15-JAN-2004 (first entry)
XX
XX Staphylococcus aureus MurC protein (SeqID 28).
XX microbial; antimicrobial; membrane biosynthesis; pathogenic;
KW immunological response; vaccination; surface disinfectant;
KW personal hygiene application; food preservative; enzyme; MurC;
KW UDP-N-acetylmuramate-alanine ligase.
XX
XX Staphylococcus aureus.
OS
XX WO2003025007-A2.
XX
XX 27-MAR-2003.
XX
XX 20-SEP-2002; 2002WO-CA001428.
XX 21-SEP-2001; 2001US-0323992P.
XX 21-SEP-2001; 2001US-0324152P.
PR 25-SEP-2001; 2001US-0324692P.
PR 26-OCT-2001; 2001US-0339924P.
PR 29-OCT-2001; 2001US-0350973P.
PR 30-OCT-2001; 2001US-0340924P.
PR 27-NOV-2001; 2001US-0333666P.
PR 18-DEC-2001; 2001US-0341732P.
PR 18-DEC-2001; 2001US-0341776P.
PR 19-DEC-2001; 2001US-0341949P.
XX
XX (AFFI-) AFFINIUM PHARM INC.
XX
XX Edwards A, Dharamsi A, Vedadi M, Alam MZ, Awrey D, Beattie B;
PI Domagala M, Houston S, Kanagarajah D, Nethery K, Ng I, Mansoury K;
PI McDonald M, Pinder B, Viola C, Wrezel O;
XX
XX WPI; 2003-468119/44.
DR N-PSDB; ADD15462.
XX
XX Novel crystallized recombinant polypeptides from Staphylococcus aureus,
PT Streptococcus pneumoniae and Escherichia coli and which are involved in

membrane biosynthesis, useful as targets for pathogenic bacteria.

Claim 20; SEQ ID NO 28; 325pp; English.

This invention relates to the structural and functional characterisation of microbial polypeptides from *Staphylococcus aureus* (S. aureus), *Streptococcus pneumoniae* (S. pneumoniae) and *Escherichia coli* (E. coli) that provide novel antimicrobial targets. Specifically, it refers to polypeptides that are involved in membrane biosynthesis, which play a critical role in the life cycle and viability of their pathogenic species of origin, and hence provide valuable drug targets. Furthermore, the invention describes modified version of these proteins that facilitate characterisation by labelling with isotopic or heavy atoms, and also fusion proteins. These proteins provide structural and functional information to aid the discovery of therapeutic molecules to treat disorders associated with a particular pathogenic species. As such, they are useful for inducing an immunological response in an individual and as an antigen for vaccination purposes. The polypeptides are also useful for developing antimicrobial agents for use as surface disinfectants, personal hygiene applications and as food preservatives or in treating food products to eliminate potential pathogens. This polypeptide sequence is the MurC protein encoded by DNA predicted from the genomic sequence of S. aureus UDP-N-acetylmuramate-alanine ligase of the invention.

Sequence 444 AA;

Query Match 100.0%; Score 2275; DB 7; Length 444;

Best Local Similarity 100.0%; Pred. No. 9.9e-193; Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTHYFVGKSGMSLAQIMHDLGHEVQSGSDIENVVTEVALRNKGKILPFDANNIKE 60
Db 8 MTHYFVGKSGMSLAQIMHDLGHEVQSGSDIENVVTEVALRNKGKILPFDANNIKE 67
Qy 61 DMVVTQGNFASHEEIVRAHQKLDVSVYNDFLGQIIDQYTSVAVTGAGKSTTGLLS 120
Db 68 DMVVTQGNFASHEEIVRAHQKLDVSVYNDFLGQIIDQYTSVAVTGAGKSTTGLLS 127
Qy 121 HVNMGDKTSLFGDTGMLPSDFAFACEYRHFSLYKPDYALMTNIDPDHDPYFK 180
Db 128 HVNMGDKTSLFGDTGMLPSDFAFACEYRHFSLYKPDYALMTNIDPDHDPYFK 187
Qy 181 DINDVDFADQEMAHNVKGGIAGWDDHLRKEADVPYIYGFKSDDIYAQNIQITDKG 240
Db 198 DINDVDFADQEMAHNVKGGIAGWDDHLRKEADVPYIYGFKSDDIYAQNIQITDKG 247
Qy 241 TAPDVYVDGEFVDHFLSPQYGDHVTNLALAVIAISYLEKLDVTNKEALETFGQVKRFN 300
Db 248 TAPDVYVDGEFVDHFLSPQYGDHVTNLALAVIAISYLEKLDVTNKEALETFGQVKRFN 307
Qy 301 ETTIANQVIVDVAHHPREISATIEATKRYPKHEVVAVFPQHTFSTQAFLEFAESLS 360
Db 308 ETTIANQVIVDVAHHPREISATIEATKRYPKHEVVAVFPQHTFSTQAFLEFAESLS 367
Qy 361 KADRVPLCEIFGSIKRGNTGALTQDLIDKIEGASLINEDSINVLEQPDNVAIFMGAGDI 420
Db 368 KADRVPLCEIFGSIKRGNTGALTQDLIDKIEGASLINEDSINVLEQPDNVAIFMGAGDI 427
Qy 421 QKLQAYLDKLGKNAF 437
Db 428 QKLQAYLDKLGKNAF 444

RESULT 5
ABU15856
ID ABU15856 standard; protein; 437 AA.
XX
AC ABU15856;

XX 19-JUN-2003 (first entry)
XX Protein encoded by Prokaryotic essential gene #1383.
DE
DE

Antisense; prokaryotic essential gene; cell proliferation; drug design.
Staphylococcus aureus.

W0200277183-A2.

03-OCT-2002.

21-MAR-2002; 2002WO-US009107.

21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00948993.

23-OCT-2001; 2001US-0342923P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI; 2003-029926/02.

N-PSDB; ACA19726.

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 25; SEQ ID NO 43780; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than S. aureus, S. typhimurium, K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 437 AA;

Query Match 100.0%; Score 2274; DB 6; Length 437;

Best Local Similarity 99.8%; Pred. No. 1.2e-192;

Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTHYFVGKSGMSLAQIMHDLGHEVQSGSDIENVVTEVALRNKGKILPFDANNIKE 60

Db 1 MTHYFVGKSGMSLAQIMHDLGHEVQSGSDIENVVTEVALRNKGKILPFDANNIKE 60

Qy 61 DMVVTQGNFASHEEIVRAHQKLDVSVYNDFLGQIIDQYTSVAVTGAGKSTTGLLS 120

61 DMVVIQGNAPASSHEEIVRAHQKLDVVSNDPLGQIIDQYTSVAVTGAHGKSTTGLLS 120
121 HVNMGDKKTSFLIGDGTGMGLPESDYPFAFEACEYRRHFLSKPDYAIMTNDIDFHPDYPFK 180
121 HVNMGDKKTSFLIGDGTGMGLPESDYPFAFEACEYRRHFLSKPDYAIMTNDIDFHPDYPFK 180
181 DINDVFAFQEMAHNVKGGIIAWGDDHRLKIEADVPYIYFGKSDSDIYAQNIQITDKG 240
181 DINDVFAFQEMAHNVKGGIIAWGDDHRLKIEADVPYIYFGKSDSDIYAQNIQITDKG 240
241 TAFDVYDGEFVDFHFLSPQYGDHTVNLALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
241 TAFDVYDGEFVDFHFLSPQYGDHTVNLALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
301 ETTIANQVIVDDYAHHPREISATITETARKKYPHEVAVFQPTFSRTQAFINEFAESLS 360
301 ETTIANQVIVDDYAHHPREISATITETARKKYPHEVAVFQPTFSRTQAFINEFAESLS 360
361 KADRVFLCEIFGSIKRENTGALTIDQIDKIEGASLINESINLVLEQFDNAVILFMGAGDI 420
361 KADRVFLCEIFGSIKRENTGALTIDQIDKIEGASLINESINLVLEQFDNAVILFMGAGDI 420
421 OKLQAYLDKLGKKNAP 437
421 OKLQAYLDKLGKKNAP 437

RESULT 6
AAU33801
ID AU33801 standard; protein; 437 AA.
XX AC AU33801;
XX AC AU33801;
DT 14-FEB-2002 (first entry)
DE DE
XX Staphylococcus aureus cellular proliferation protein #77.
XX Staphylococcus aureus cellular proliferation protein; antibiotic;
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX antibacterial; drug design.
XX Staphylococcus aureus.
OS OS
XX WO200170955-A2.
XX WO200170955-A2.
XX 27-SEP-2001.
XX 21-MAR-2001; 2001WO-US009180.
XX 21-MAR-2000; 2000US-0191078P.
XX 23-MAY-2000; 2000US-0206848P.
XX 26-MAY-2000; 2000US-0207272P.
XX 23-OCT-2000; 2000US-0242578P.
XX 27-NOV-2000; 2000US-0253625P.
XX 22-DEC-2000; 2000US-0257931P.
XX 16-FEB-2001; 2001US-0269308P.
XX (ELIT-) ELITRA PHARM INC.
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX WPI: 2001-611495/70.
XX N-PSDB; AAS1660.
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids.
XX Example 3; SEQ ID NO 5297; 511pp; English.
XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the genes,
XX their use in the discovery of novel antibiotics, the essential genes

CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 437 AA;
SQ

Query Match 99.6%; Score 2267; DB 4; Length 437;
Best Local Similarity 99.5%; Pred. No. 5e-192;
Matches 435; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTHYHFVGIKSGMSLSAQIMHDLGHEVQSGSDIENVVTEVALRNKGIKILPPDANNIKE 60
DB 1 MTHYHFVGIKSGMSLSAQIMHDLGHEVQSGSDIENVVTEVALRNKGIKILPPDANNIKE 60
QY 61 DMVVIQGNAPASSHEEIVRAHQKLDVVSNDPLGQIIDQYTSVAVTGAHGKSTTGLLS 120
DB 61 DMVVIQGNAPASSHEEIVRAHQKLDVVSNDPLGQIIDQYTSVAVTGAHGKSTTGLLS 120
QY 121 HVNMGDKKTSFLIGDGTGMGLPESDYPFAFEACEYRRHFLSKPDYAIMTNDIDFHPDYPFK 180
DB 121 HVNMGDKKTSFLIGDGTGMGLPESDYPFAFEACEYRRHFLSKPDYAIMTNDIDFHPDYPFK 180
QY 181 DINDVFAFQEMAHNVKGGIIAWGDDHRLKIEADVPYIYFGKSDSDIYAQNIQITDKG 240
DB 181 DINDVFAFQEMAHNVKGGIIAWGDDHRLKIEADVPYIYFGKSDSDIYAQNIQITDKG 240
QY 241 TAFDVYDGEFVDFHFLSPQYGDHTVNLALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
DB 241 TAFDVYDGEFVDFHFLSPQYGDHTVNLALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
QY 301 ETTIANQVIVDDYAHHPREISATITETARKKYPHEVAVFQPTFSRTQAFINEFAESLS 360
DB 301 ETTIANQVIVDDYAHHPREISATITETARKKYPHEVAVFQPTFSRTQAFINEFAESLS 360
QY 361 KADRVFLCEIFGSIKRENTGALTIDQIDKIEGASLINESINLVLEQFDNAVILFMGAGDI 420
DB 361 KADRVFLCEIFGSIKRENTGALTIDQIDKIEGASLINESINLVLEQFDNAVILFMGAGDI 420
QY 421 OKLQAYLDKLGKKNAP 437
DB 421 OKLQAYLDKLGKKNAP 437

RESULT 7
AAU36700
ID AU36700 standard; protein; 444 AA.
XX AC AU36700;
XX AC AU36700;
DT 14-FEB-2002 (first entry)
DE DE
XX Staphylococcus aureus cellular proliferation protein #870.
XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX antibacterial; drug design.
XX Staphylococcus aureus.
XX WO200170955-A2.
XX 27-SEP-2001.
PD

XX PF 21-MAR-2001; 2001WO-US009180.
XX PR 21-MAR-2000; 2000US-0191078P.
XX PR 23-MAY-2000; 2000US-0206848P.
XX PR 26-MAY-2000; 2000US-0207272P.
XX PR 23-OCT-2000; 2000US-0242578P.
XX PR 27-NOV-2000; 2000US-0253625P.
XX PR 22-DEC-2000; 2000US-0257931P.
XX PR 16-FEB-2001; 2001US-0269308P.
XX FA (ELIT-) ELITRA PHARM INC.
XX XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yamamoto RT, Xu HH;
XX XX WPI; 2001-611495/70.
XX DR N-PSDB; AAS54559.
XX DR New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids.
XX PT
XX ES Example 3; SEQ ID NO 12293; 511pp; English.
XX XX
XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the genes,
XX CC their use in the discovery of novel antibiotics, the essential genes
XX CC themselves and the encoded proteins. The prokaryotes used are Escherichia
XX CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX CC useful for the identification of potential new targets for antibiotic
XX CC development. The antisense nucleic acids can also be used to identify
XX CC proteins used in proliferation, to express these proteins, and to obtain
XX CC antibodies capable of binding to the expressed proteins. The proteins can
XX CC be used to screen compounds in rational drug discovery programmes. The
XX CC antisense nucleic acid sequence is also useful to screen for homologous
XX CC nucleic acids which are required for cell proliferation in a wide variety
XX CC of organisms. The present sequence represents an essential prokaryotic
XX CC cellular proliferation protein. Note: The sequence data for this patent
XX CC did not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX XX
XX XX Sequence 444 AA;
XX XX
XX XX Query Match 99.6%; Score 2267; DB 4; Length 444;
XX XX Best Local Similarity 99.5%; Pred. No. 5.1e-192;
XX XX Matches 435; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX XX
XX QY 1 MTHYFVGIKSGMSSLAQIMHDLGHEVQGSIDENYVTFEVALRNKGIKILPFDANNIKE 60
XX DB 8 MTHYFVGIKSGMSSLAQIMHDLGHEVQGSIDENYVTFEVALRNKGIKILPFDANNIKE 67
XX QY 61 DMVVIQGNAFASHEEIVRAHQIKLDVSYNDFLGQIIDQYTSVAVTGAKGTSTTGLLS 120
XX DB 68 DMVVIQGNAFASHEEIVRAHQIKLDVSYNDFLGQIIDQYTSVAVTGAKGTSTTGLLS 127
XX QY 121 HVMNGDKTSTFLIGDGTGMLPSDFAFACEYRRHFLSYKPDYAIMTNIDPDHEDYFK 180
XX DB 128 HVMNGDKTSTFLIGDGTGMLPSDFAFACEYRRHFLSYKPDYAIMTNIDPDHEDYFK 187
XX QY 181 DINDVDFAFQEAHNVKGIIGWDDHRLKTEADVPVYVYKSDSDDIYAQNIQITDKG 240
XX DB 188 DINDVDFAFQEAHNVKGIIGWDDHRLKTEADVPVYVYKSDSDDIYAQNIQITDKG 247
XX QY 241 TAFDVYVDGEFVDHFLSPQYGDHTVNLALAVIAISYLEKLDVTNIKEALETFGKVRREN 300
XX DB 248 TAFDVYVDGEFVDHFLSPQYGDHTVNLALAVIAISYLEKLDVTNIKEALETFGKVRREN 307
XX QY 301 ETTIANQIVDDYAHHPREISATETARXYPKHEVAVFPQHTFRTQAFINFEAELS 360
XX DB 308 ETTIANQIVDDYAHHPREISATETARXYPKHEVAVFPQHTFRTQAFINFEAELS 367

QY 361 KADRVFLCEIFGSIRESNTGALTITQDLIDKLEGASLINESINVLEQFDNAVILFMGAGDI 420
DB 368 KADRVFLCEIFGSIRESNTGALTITQDLIDKLEGASLINESINVLEQFDNAVILFMGAGDI 427
QY 421 QKLNAYLDKLGKMNAF 437
DB 428 QKLNAYLDKLGKMNAF 444
RESULT 8
AAW87771
ID AAW87771 standard; protein; 437 AA.
XX AAW87771;
XX AC AAW87771;
XX DT 17-MAR-1999 (first entry)
XX XX UDP-N-acetylmuramate:L-alanine ligase (MurC polypeptide).
XX DE
XX XX MurC gene; UDP-N-acetylmuramate:L-alanine ligase; MurC polypeptide;
XX KW bacterial; infection; H. pylori; cancer; ulcer; gastritis; vaccine;
XX KW immunogen; drug; genetic immunisation.
XX XX Staphylococcus aureus.
XX OS
XX PN EP889123-A2.
XX XX
XX PD 07-JAN-1999.
XX XX
XX PF 26-JUN-1998; 98EP-00305064.
XX XX
XX PR 03-JUL-1997; 97US-0052720P.
XX XX (SMIK) SMITHKLINE BEECHAM CORP.
XX PA (SMIK) SMITHKLINE BEECHAM PLC.
XX XX
XX PI Wallis NG, Burnham MKR;
XX XX WPI; 1999-062655/06.
XX DR N-PSDB; AAV99650.
XX XX
XX PT New isolated MurC polypeptide from Staphylococcus aureus and related
XX PT nucleic acid - useful in diagnosis, treatment and prevention of bacterial
XX PT infections.
XX XX
XX PS Claim 1; Page 4; 39pp; English.
XX XX
XX CC The present sequence represents a UDP-N-acetylmuramate:L-alanine ligase
XX CC (MurC polypeptide) encoded by the S. aureus MurC gene. Host cells
XX CC containing an expression system comprising the MurC gene can be used for
XX CC the recombinant production of the polypeptide. Agonists or the MurC
XX CC polypeptide are used to treat conditions requiring increased activity or
XX CC expression of the polypeptide. Antagonists, inhibitory nucleic acid or
XX CC competitive polypeptide are useful for inhibiting the polypeptide e.g.
XX CC bacterial (especially S. aureus) infections. They are also useful against
XX CC Helicobacter pylori infections and related cancers, ulcers and gastritis.
XX CC The antibacterial agents are useful to treat in-dwelling devices for
XX CC infection prevention or generally as wound treatments to prevent adhesion
XX CC of bacteria to matrix proteins. The MurC polypeptide is also useful for
XX CC diagnosing or prognosing a (susceptibility to) disease, for raising
XX CC antibodies; to identify modulators or specific receptors; in rational
XX CC drug design and as an immunogen for vaccines. The MurC gene sequences are
XX CC useful in antisense/ribozyme therapeutics; to detect mutant MurC gene;
XX CC for chromosomal mapping; to determine bacterial serotype; and for genetic
XX CC immunisation
XX XX
XX SQ Sequence 437 AA;
XX XX
XX XX Query Match 98.4%; Score 2238; DB 2; Length 437;
XX XX Best Local Similarity 98.2%; Pred. No. 1.8e-189;
XX XX Matches 429; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
XX XX
XX QY 1 MTHYFVGIKSGMSSLAQIMHDLGHEVQGSIDENYVTFEVALRNKGIKILPFDANNIKE 60

1 MTHYHFVGKSGMSSLAQIMHDLGHEVQSGSDIENYVTFEVALRNKGKILPFGANNIK 60
61 DMVVIQGNAPASSHEEIVRAHQKLDVVSNDVFLGQIIDQVTSVAVTGAGHKTSTTGLS 120
61 DMVVIQGNAPASSHEEIVRAHQKLDVVSNDVFLGQIIDQVTSVAVTGAGHKTSTTGLS 120
121 HVNMGDKKTSFLIGDGTGMLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDPDYFK 180
121 HVNMGDKKTSFLIGDGTGMLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDPDYFK 180
181 DINDVDFAFQEMAHNVKGGIIAWGDDDEHLRKIEADVPIYYGFKDSDDIYAQNIQITDKG 240
181 DINDVDFAFQEMAHNVKGGIIAWGDDDEHLRKIEADVPIYYGFKDSDDIYAQNIQITDKG 240
241 TAFDVVVDGFEYDFHFLSPQYGDHVTMLNALAVIAISYLEKLDVTNIKEALETFGGVKRNF 300
241 TAFDVVVDGFEYDFHFLSPQYGDHVTMLNALAVIAISYLEKLDVTNIKEALETFGGVKRNF 300
301 ETTIANQVIVDDYAHHPREISATITETARKKYPHKEVVAVFQPHTFSTRTOAFLEFAESLS 360
301 ETTIANQVIVDDYAHHPREISATITETARKKYPHKEVVAVFQPHTFSTRTOAFLEFAESLS 360
361 KADRVFLCEIFGSIRESNTGALTQDLIDKIEGASLINESINVLQFDNAVILFMGAGDI 420
361 KADRVFLCEIFGSIRESNTGALTQDLIDKIEGASLINESINVLQFDNAVILFMGAGDI 420
421 OKLQNAVYLDKLGKNAF 437
421 OKLQNAVYLDKLGKNAF 437

RESULT 9
ABP40288
ID ABP40288 standard; protein; 442 AA.
XX
XX
XX
24-JUL-2002 (first entry)
XX
XX
Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5133.
XX
XX
Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
XX
XX
antibacterial; gene therapy.
XX
XX
Staphylococcus epidermidis.
XX
XX
US6380370-B1.
XX
XX
30-APR-2002.
XX
XX
13-AUG-1998; 98US-00134001.
XX
XX
14-AUG-1997; 97US-0055779P.
XX
XX
08-NOV-1997; 97US-0064964P.
XX
XX
(GENO-) GENOME THERAPEUTICS CORP.
XX
XX
Doucette-Stamm LA, Bush D;
XX
XX
WPI; 2002-381255/41.
XX
XX
N-PSDB; ABN92833.
XX
XX
Novel isolated nucleic acid encoding a Staphylococcus epidermis
XX
XX
polypeptide, useful for diagnosing and treating bacterial infections.
XX
XX
Disclousure; SEQ ID NO 5133; 267pp; English.
XX
XX
ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
XX
XX
frame (ORF) nucleic acid sequences which encode the amino acid sequences
XX
XX
given in ABP35124 to ABP37960. The S. epidermidis sequences have
XX
XX
antibacterial activity and can be used in gene therapy. The sequences can
XX
XX
also be used in the diagnosis and treatment of bacterial infections,
XX

CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life cycle
CC or inhibit S. epidermidis infection. N.B. The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the USPTO web site
XX

SQ Sequence 442 AA;

Query Match 87.3%; Score 1986; DB 5; Length 442;
Best Local Similarity 84.7%; Pred. No. 4e-167;
Matches 370; Conservative 43; Mismatches 24; Indels 0; Gaps 0;
QY 1 MTHYHFVGKSGMSSLAQIMHDLGHEVQSGSDIENYVTFEVALRNKGKILPFGANNIK 60
DB 6 MTHYHFVGKSGMSSLAQIMHDLGHEVQSGSDIENYVTFEVALRNKGKILPFGANNITK 65
QY 61 DMVVIQGNAPASSHEEIVRAHQKLDVVSNDVFLGQIIDQVTSVAVTGAGHKTSTTGLS 120
DB 66 EMVVIQGNAPFPDNHHEIVRAHELKLDIIKIYHDFLGHVINOYTSVAVTGAGHKTSTTGLS 125
QY 121 HVNMGDKKTSFLIGDGTGMLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDPDYFK 180
DB 126 HVNMGDKKTSFLIGDGTGMLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDPDYFK 185
QY 181 DINDVDFAFQEMAHNVKGGIIAWGDDDEHLRKIEADVPIYYGFKDSDDIYAQNIQITDKG 240
DB 186 NIDVDYDAFQEMAHNVKGGIIAWGDDDEHLRKIEADVPIYYGFKDSDDIYAQNIQITDKG 245
QY 241 TAFDVVVDGFEYDFHFLSPQYGDHVTMLNALAVIAISYLEKLDVTNIKEALETFGGVKRNF 300
DB 246 TQFDVYIKGEFYDFHFLSPQYGNHILNALAVIAISYLENNVNIKEALITFGGVKRRFN 305
QY 301 ETTIANQVIVDDYAHHPREISATITETARKKYPHKEVVAVFQPHTFSTRTOAFLEFAESLS 360
DB 306 ETKVSNQVIVDDYAHHPREISATITETARKKYPQKQVAVFQPHTFSTRTOAFLEFAESLS 365
QY 361 KADRVFLCEIFGSIRESNTGALTQDLIDKIEGASLINESINVLQFDNAVILFMGAGDI 420
DB 366 KADQVFLCEIFGSIRESNTGALTIEDLINRIGDSTLIDENSIDVLEKFDNAVILFMGAGDI 425
QY 421 OKLQNAVYLDKLGKNAF 437
DB 426 OKLLKAYFEKLGKND 442

RESULT 10
ABU43254
ID ABU43254 standard; protein; 437 AA.
XX
XX
AC ABU43254;
XX
XX
19-JUN-2003 (first entry)
XX
XX
Protein encoded by Prokaryotic essential gene #28781.
XX
XX
Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX
Staphylococcus epidermidis.
XX
XX
WO200277183-A2.
XX
XX
PD 03-OCT-2002.
XX
XX
21-MAR-2002; 2002WO-US009107.
XX
XX
21-MAR-2001; 2001US-00815242.
XX
XX
06-SEP-2001; 2001US-00948993.
XX
XX
25-OCT-2001; 2001US-0342923P.
XX
XX
08-FEB-2002; 2002US-00072851.
XX
XX
06-MAR-2002; 2002US-0362699P.
XX
XX
(ELIT-) ELITRA PHARM INC.
XX

QY 304 IANQVIVDDYAHHPREISATETARKKYPHKEVAVFQPHFTSRTQAFPLNEFAESLSKAD 363
 Db VSDMIIVDDYAHHPAEIKATIDGARQKYPDKKEIIAVFQPHFTTTRTALMDEFAELDLAD 368
 QY 364 RVFLCEIFGSIRENTGALTIQDLIDKIE-GASLINEDSINVLEQEDNAVILFMGAGDIQK 422
 Db 369 EVFLCNIFGSARETOGEVRIEDLGEKIQKGGQVITEDNVSLDDEFNAVVFVGAGDVQK 428
 QY 423 LONAY 427
 Db 429 FEQAY 433

AAU35277
 ID AAU35277 standard; protein; 456 AA.
 AC AAU35277;
 DT 14-FEB-2002 (first entry)
 DE Enterococcus faecalis cellular proliferation protein #564.
 XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 OS Enterococcus faecalis.
 PN WO200170955-A2.
 XX 27-SEP-2001.
 PF 21-MAR-2001; 2001WO-US009180.
 XX 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX (ELIT-) ELITRA PHARM INC.
 PA Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 PI WPI; 2001-611495/70.
 DR N-PSDB; AA53136.
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 XX Example 3; SEQ ID NO 10870; 51pp; English.
 XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes,
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 456 AA;
 Query Match 63.0%; Score 1433.5; DB 4; Length 456;
 Best Local Similarity 60.7%; Pred. No. 3.9e-118;
 Matches 258; Conservative 79; Mismatches 87; Indels 1; Gaps 1;
 QY 4 YHFVGIKSGMSSLAQIMHDLGHEVQSGSDIENYVFEVALRNKGKIKLPEDANNIKEDMV 63
 Db 20 YHFVGIKSGMSSLAQIMHDLGHEVQSGSDIENYVFEVALRNKGKIKLPEDANNIKEDMV 79
 QY 64 VIQNAFASHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSITGLLSHYM 123
 Db 80 VIAGNAPDTHETARAEIQAERIRYHDFIARFIEPTTSIAVTSKGKTSITGLLAHVL 139
 QY 124 NGDKTSLFLGDTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDPKDN 183
 Db 140 SGINPTSLVIGDTGHEGPDADFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDPKDN 199
 QY 184 DVFAFQEMAHNVKGIITANGDDEHLRLKLEADVPYIYVYGFKSDDIYAQNIQITDKGTAF 243
 Db 200 DVFSAFQTMHQVKKIPAYGDKYLRQLESEVPYIYVYGVSEEDDIQAENIORTTEGSSF 259
 QY 244 DVYVDGEYDFHFLSPQYGDHTVLAALAVIAISYLEKLDVTNIKEALETEGGVKKRNETT 303
 Db 260 DVYHKADPVGHFVLPFAGHHNMNMGVIAVAFKLDQKVAEMLSEFGVKRRFSEKK 319
 QY 304 IANQVIVDDYAHHPREISATETARKKYPHKEVAVFQPHFTSRTQAFPLNEFAESLSKAD 363
 Db 320 VSDMIIVDDYAHHPAEIKATIDGARQKYPDKKEIIAVFQPHFTTTRTALMDEFAELDLAD 379
 QY 364 RVFLCEIFGSIRENTGALTIQDLIDKIE-GASLINEDSINVLEQEDNAVILFMGAGDIQK 422
 Db 380 EVFLCNIFGSARETOGEVRIEDLGEKIQKGGQVITEDNVSLDDEFNAVVFVGAGDVQK 439
 QY 423 LONAY 427
 Db 440 FEQAY 444

RESULT 14
 ABU29727
 ID ABU29727 standard; protein; 444 AA.
 XX ABU29727;
 AC ABU29727;
 DT 19-JUN-2003 (first entry)
 DE Protein encoded by Prokaryotic essential gene #15254.
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW Enterococcus faecium.
 OS Enterococcus faecium.
 XX WO200277183-A2.
 XX 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA33597.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 57651; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 444 AA;

Query Match 61.4%; Score 1396; DB 6; Length 444;
Best Local Similarity 60.6%; Pred. No. 7.9e-115;
Matches 260; Conservative 74; Mismatches 93; Indels 2; Gaps 2;

QY 4 YHFGIKSGMSSLAQIMHDLGHEVQSGSDIENYVTEVALRNKGILKILPFDANNIKEDMV 63
DB 8 YHFGIKSGMSSLAQIMHDLGHEVQSGSDIENYVTEVALRNKGILKILPFDANNIKPGMT 67
QY 64 VIQGNFASHEEIVRAHQLKLDVVSYNDFLGOIIDQVTSVAVTGAGHKTSTTGLLSHMV 123
DB 68 ILAGNAPDSHEEIQRAKELGLEVIRYHDFGHFTQNTYSIAVTSHGKTSITGLLSHVL 127
QY 124 NGDKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSKYKPDYVIMTNDIDFDPYFKDIN 183
DB 128 SGVRPTSYLIGDGTGHGDPQAEFFSFEACEYRRHFLAYSPPVIMTNDIDFDPYVTSID 187
QY 184 DVFPDFQEMAHNVKGIITANGDDEHLKLEADVPYIYVGFKSDDIYAQNIQITDKGTAP 243
DB 188 DVFTAFQTMAGQVKAIFAYGSDAYRLKANKVPYIYVGVNTDNDIOARNIERTSGSAP 247
QY 244 DVYVGGEPDYHFLSPQYGDHTVNLALVIAISYLEKLDVTNKEALETFGGVKRRFNETT 303
DB 248 DVYHGDVFGHTVPAFGKHNILNALGVIAVAYFKLDLKEVAEMLTTPPGVKRRFSEKI 307
QY 304 IANQVIVDDYAHHPREISATITETARKKYPHKEVAVFQPHFSTRQAFINEFAESLSKAD 363
DB 308 VADMTVVDDYAHHPAEIKATIDGARQKYPDKREIIAVFQPHFTRTIALMDEFAEALDLAD 367
QY 364 RVFICEITFGSTRENTGALTIOQLDKI-EGASLINESINVLQEPDNVILPMGAGDIQK 422
DB 368 KVYLCDFIGSAREEQNVKIEDLGAKIKKGGEVINKENNVSPLLDYDAVVIWMGAGDVQK 427

QY 423 LQNAVYLDKL 431
DB 428 FEQAY-EKL 435

RESULT 15
ABB48762

ID ABB48762 standard; protein; 447 AA.

XX ABB48762;

DT 05-FEB-2002 (first entry)

DE Listeria monocytogenes protein #1465.

XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.

XX Listeria monocytogenes.

PN WO200177335-A2.

PD 18-OCT-2001.

PF 11-APR-2001; 2001WO-FR001118.

PR 11-APR-2000; 2000FR-00004629.

XX (INSP) INST PASTEUR.

XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetoui F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kneft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX WPI; 2002-010914/01.

XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and related
PT polypeptides.

XX Claim 6; SEQ ID NO 1467; 192pp; French.

XX The present invention relates to the genome sequence of Listeria
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in L.
CC monocytogenes and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present sequence is a protein
CC encoded by the genome sequence of the present invention. Proteins
CC expressed from the genome sequence are useful for raising specific
CC antibodies, identification of L. monocytogenes and related organisms, and
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate L. monocytogenes-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and
CC vaccine compositions for the treatment or prevention of infections by L.
CC monocytogenes and related organisms. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 447 AA;

Query Match 59.0%; Score 1342; DB 5; Length 447;
Best Local Similarity 59.7%; Pred. No. 4.9e-110;
Matches 255; Conservative 64; Mismatches 108; Indels 0; Gaps 0;

QY 1 MTHYHFVGKSGMSSLAQIMHDLGHEVQSGSDIENYVTEVALRNKGILKILPFDANNIKE 60

```
Db      1  MTIYHVGKSGNSALAQILHDKGFQVQSGSDVKYFTQKALEBKQIPIMTFSSADNIQE 60
QY      61  DMVVIQNAFASHEEIVRAHQKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
Db      61  GLATIAGNAPDTHTEIERALELGLSVIRYHKFLGQLIDGVTSAITGSHGKTSTTGLLS 120
QY      121  HVNMGDKKTSFLIGDGTGMLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
Db      121  HVGAIKPTSYLIGDGTGSGTKDAEYFALEACEYQRHFLAYKPTYAIMTNIDWDHPDYFK 180
QY      181  DINDVDAFOEMAHNVKGIILAWGDDEHLRXIEADVPYIYYGFKSDDDIYAQNIQITDKG 240
Db      181  SVDDVFNAFETLKGQVKKAVFALGDDAELEKRLTLDIPYIFGGEENEFOAKNVIKETTG 240
QY      241  TAFDVYVDGEFYDHFSLSPQYGDHTVINALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
Db      241  TKFDVYHREEFLLSFEIPAYGDHVNALSVIALCDYEGLPVEDVKNELKTFEGVKRFFS 300
QY      301  ETTIANQVIVDDYAHHPREISATITETARKKYPHKEVVAVFOPHTFSRTQAFINEFAESLS 360
Db      301  ITEKGNQVLYDDYAHHPSEIRATVNAARQKYPDKKVVAVFOPHTFTRTRTELQGFADSLN 360
QY      361  KADRVFLCEIFGSLRENTGALTIODLIDKIEGASLINEDSINVLEQPDNAVILFMAGDI 420
Db      361  LADEVILCDIFGSAKEXTGNLTIAHLAKTKGNHIIKEEHTPELLKYPEAVILFMAGDV 420
QY      421  QKLQNAV 427
Db      421  QKFOAAY 427
```

Search completed: June 3, 2004, 14:39:09
Job time : 61 secs

QY 361 KADRVFLCEIFGSIKRENTGALTIQDLIDKIEGASLINESINVLEQFDNAVILFMGAGDI 420
Db 366 KADQVFLCEIFGSIKRENTGALTIQDLIDKIEGASLINESINVLEQFDNAVILFMGAGDI 425
QY 421 OKLQNAVILKMGKNAF 437
Db 426 OKLLKAYFEKLGVRKND 442

RESULT 2

US-08-940-572-2
; Sequence 2, Application US/08940572
; Patent No. 6310193
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; APPLICANT: Black, Michael T.
; APPLICANT: Hodgson, John E.
; APPLICANT: Knowles, David J.
; APPLICANT: Lonetto, Michael A.
; APPLICANT: Nicholas, Richard O.
; APPLICANT: Stodola, Robert K.
; TITLE OF INVENTION: No. 6310193el Murc
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/940,572
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024022
; FILING DATE: 16-AUG-1996
; APPLICATION NUMBER: US 08/889711
; FILING DATE: 08-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Q. Todd
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50533-04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215/994-2252
; TELEFAX: 215/994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 444 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-940-572-2

Query Match 49.2%; Score 1119; DB 4; Length 444;
Best Local Similarity 49.8%; Pred. No. 7e-101;
Matches 213; Conservative 84; Mismatches 127; Indels 4; Gaps 4;
QY 4 YHFGVIGKSGMSLAQIMHDLGHEVQSGSDIENVVTEVALRNKGIKILPFDANNIKEDMV 63
Db 5 YHFGVIGKSGMSALMLHQMKGKVGQSDVEKYFFQGLEQAGITILPFDENKLDGME 64
QY 64 VTQGNF-ASSHEEIVRAHQKLDVVSNDPLGQIIDQYTSVAVTGAGKGTSTTGLSHV 122
Db 65 LIAGNAFPDNNVEIAYAQNGISYKRYHEFLGSEMRDFVSMGVAGNKGKTSITGMLSHV 124
QY 123 MNGDKKTSFLIGDGTGNGLPESDYFAFEACEYRRHFLSYKPDYAIMTNTIDFDHPDYFKOI 182

Db 125 LSHITTSFLIGDGTGRGSANAKYVFESDEYERHFMPEYIITNIDFDHPDYFTSL 184
QY 183 NDVFADFOEMAHNVKKGIIIAWGDDDEHLRKEADVPYIYVYGFK-DSDDIYAQNIQITDKGT 241
Db 185 EDVENAFNDYAKOITKGLFVYGDAELRKITSDAPIYVYGEAGNDVFASDLRSTTGS 244
QY 242 AFDVYDGEFYDHPHLSPOYGDHVTNLALAVIAISYLEKLDVTNIKEALETFGGVKRRFNE 301
Db 245 TFTVHFRGQNLGQFHPTFRGHNMNATAVIGLITAGFDLNLVREHLKTFAGVKRRFTE 304
QY 302 TTIANOVIVDDYAHHPREISATITARKKYPHKEVAVFQPHITSRTQAFNLNEFAESLSK 361
Db 305 KIVNDTVIIDDFAHPTETIATLDAARQKYPSEIIVAVFQPHITRTTIALDDFAHALNQ 364
QY 362 ADRVFLCEIFGSIKRE-NTGALTITQDLIDKI-EGASLINESINVLEQFDNAVILFMGAGD 419
Db 365 ADAVYLAQIYGSAREYVDHGDVKVEDLANLKKHQVITVENVSELLDHDNAVYVFMGAGD 424
QY 420 IQKLQNAV 427
Db 425 IQTYEYSF 432

RESULT 3

US-08-961-083-116
; Sequence 116, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 116:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 422 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-083-116

Query Match 45.8%; Score 1043; DB 3; Length 422;
Best Local Similarity 48.8%; Pred. No. 1.7e-93;
Matches 200; Conservative 80; Mismatches 126; Indels 4; Gaps 4;
QY 22 HDLGHEVQSGSDIENVVTEVALRNKGIKILPFDANNIKEDMVVIQGNF-ASSHEEIVRA 80

Db 1 HQMGHKVQGSDEVEKYFTQGLEQAGITILPFDEKNDLGDMEIAGNAPRDPNNVEIAYA 60
QY 81 HQKLDVVSNDPLGQIIDQYTSVAVTGAHGTSTTGLLSHVMMGDKKTSFLIGDGTGMG 140
Db 61 DQNGISYKRYHEFLGSMRDFVSMGVAGAGKSTTGMLSHVLSHITDTSFLIGDGTGRG 120
QY 141 LPSDYAEACEYERHFLSYKPDYAIMNIDPDHEDYFKDINDVDFAPQEMAHNVKGI 200
Db 121 SANAKYFVESDEYERHFMPEYHPEYSIITNIDPDHEDYFSTLSDVFNFAFYAKQITKGL 180
QY 201 IAMGDDEHLRKIEADVPYIYGYGFK-DSDDIYAQNIQITDKGTAFDVVVOGEFYDHFSLSPQ 259
Db 181 FVYGEDAELRKITSADAPIYIYGYGEAGNDVFASDLRSITGSTFTVHFRQNLGQHPIPT 240
QY 260 YGDHTVNLALAVIAISYLEKLDVNIKEALETFGGVKRRPNETTIANQVIVDDYAHHPRE 319
Db 241 FGRHNMNATAVIGLLYTAGFDLNLVREHLKTFAGVKRRFTEKIVNDTVIIDDFAHHPTE 300
QY 320 ISATLETARKYPKHEVAVFOPHTSRCTQAFNEFAESLSKADRVFLCEIFGSIRE-NT 378
Db 301 IATLDAARQKYPKSEIVAVFOPHTFTRTIALDDFAHALNQAADAVYLAQIYGSAREVDH 360
QY 379 GALTIQDLIDKI-EGASLINESINVLQFONAVILFMGAGDIOKLQNAV 427
Db 361 GDVKVEDLANKINKKHQVITVENVSPLLDHDAVYVFMGAGDIQTYEYSF 410

RESULT 4

US-09-536-784-116
; Sequence 116, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/536,784

FILING DATE: 30-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/961,083

FILING DATE: OCT-30-1997

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB340P3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 116:

SEQUENCE CHARACTERISTICS:

LENGTH: 422 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 116:

Query Match 45.8%; Score 1043; DB 4; Length 422;
Best Local Similarity 48.8%; Pred. No. 1.7e-93;

Matches 200; Conservative 80; Mismatches 126; Indels 4; Gaps 4;
QY 22 HDLGEVQGSDEIENVFTEVALRNKGIKILPDPANNIKEDMVMVIOGNAP-ASSHEIIVRA 80
Db 1 HQMGHKVQGSDEVEKYFTQGLEQAGITILPFDEKNDLGDMEIAGNAPRDPNNVEIAYA 60
QY 81 HQKLDVVSNDPLGQIIDQYTSVAVTGAHGTSTTGLLSHVMMGDKKTSFLIGDGTGMG 140
Db 61 DQNGISYKRYHEFLGSMRDFVSMGVAGAGKSTTGMLSHVLSHITDTSFLIGDGTGRG 120
QY 141 LPSDYAEACEYERHFLSYKPDYAIMNIDPDHEDYFKDINDVDFAPQEMAHNVKGI 200
Db 121 SANAKYFVESDEYERHFMPEYHPEYSIITNIDPDHEDYFSTLSDVFNFAFYAKQITKGL 180
QY 201 IAMGDDEHLRKIEADVPYIYGYGFK-DSDDIYAQNIQITDKGTAFDVVVOGEFYDHFSLSPQ 259
Db 181 FVYGEDAELRKITSADAPIYIYGYGEAGNDVFASDLRSITGSTFTVHFRQNLGQHPIPT 240
QY 260 YGDHTVNLALAVIAISYLEKLDVNIKEALETFGGVKRRPNETTIANQVIVDDYAHHPRE 319
Db 241 FGRHNMNATAVIGLLYTAGFDLNLVREHLKTFAGVKRRFTEKIVNDTVIIDDFAHHPTE 300
QY 320 ISATLETARKYPKHEVAVFOPHTSRCTQAFNEFAESLSKADRVFLCEIFGSIRE-NT 378
Db 301 IATLDAARQKYPKSEIVAVFOPHTFTRTIALDDFAHALNQAADAVYLAQIYGSAREVDH 360
QY 379 GALTIQDLIDKI-EGASLINESINVLQFONAVILFMGAGDIOKLQNAV 427
Db 361 GDVKVEDLANKINKKHQVITVENVSPLLDHDAVYVFMGAGDIQTYEYSF 410

RESULT 5

US-09-107-532A-6914
; Sequence 6914, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781) 893-5007

TELEFAX: (781) 893-8277

INFORMATION FOR SEQ ID NO: 6914:

SEQUENCE CHARACTERISTICS:

LENGTH: 291 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...291
SEQUENCE DESCRIPTION: SEQ ID NO: 6914:
US-09-107-532A-6914

Query Match 42.5%; Score 967; DB 4; Length 291;
Best Local Similarity 63.8%; Pred. No. 2.7e-86;
Matches 178; Conservative 43; Mismatches 58; Indels 0; Gaps 0;

QY 4 YHFGVIGKSGMSLAQIMHDLGHEVQSGDIENYVTFEVALRNKGIKILPFDANNIKEDMV 63
DB 13 YHFGVIGKSGMSLAQIMHDLGHEVQSGDIENYVTFEVALRNKGIKILPFDANNIKEDMV 72
QY 64 VIOGNAPASSHEEIVRAHQKLDVSYNDFLQIIDDQYTSVAVTGAHGTSTTGLLSHYM 123
DB 73 ILAGNAPDSHEEIQRAKELGLEIVIRYHDFIGHFIQNYTSIAVTGSHGKTSTTGLLSHYL 132
QY 124 NGDKTSFLIGDTGMLPESDYFAFEACEYRHRHFLSYKPDYAIMTNIDPDHDPKIN 183
DB 133 SGVPTSYLIGDTGMLPESDYFAFEACEYRHRHFLSYKPDYAIMTNIDPDHDPKIN 192
QY 184 DVYDFQEMAHNVKGIILAWGDDDEHLRKEADVPYIYVYFGKSDDDIYAQNIQITDKGTAF 243
DB 193 DVYDFQEMAHNVKGIILAWGDDDEHLRKEADVPYIYVYFGKSDDDIYAQNIQITDKGTAF 252
QY 244 DVYDFQEMAHNVKGIILAWGDDDEHLRKEADVPYIYVYFGKSDDDIYAQNIQITDKGTAF 282
DB 253 DVYDFQEMAHNVKGIILAWGDDDEHLRKEADVPYIYVYFGKSDDDIYAQNIQITDKGTAF 291

RESULT 6

US-09-328-352-5023
; Sequence 5023, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5023
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5023

Query Match 25.2%; Score 574; DB 4; Length 492;
Best Local Similarity 34.1%; Pred. No. 1.7e-47;
Matches 154; Conservative 86; Mismatches 171; Indels 40; Gaps 14;

QY 3 HYFVGIGKSGMSLAQIMHDLGHEVQSGDIENYVTFEVALRNKGIKI-LPFDANNIKED 61
DB 34 HIFVGIGKSGMSLAQIMHDLGHEVQSGDIENYVTFEVALRNKGIKI-LPFDANNIKED 92
QY 62 MVVIOGNAPASSHEEIVRAHQKLDVSYNDFLQIIDDQYTSVAVTGAHGTSTTGLLSH 121
DB 93 NVLWVSTADIPENPEVKAALQRIPIVRRABMLGELMYRHRHFIAGVAGHGTSTTGLLSH 152
QY 122 VMNGDK-KTSFLIG-----DGTGMLPESDYFAFEACEYRHRHFLSYKPDYAIMTNIDPDH 175
DB 153 MIAEENLDPTVYVIGLLNSGVNRAALGESRFVAEADSDASFLYLPMAIVTNIDADH 212
QY 176 PD-YFKDINDVDFQEMAHNVK-KGI-IAWGDDDEHLRKEADVPYIYVYFGKSDDDIY 230
DB 213 MDYEGSFDKLDKDTFQVFLHNLPPYGLAVGCGDDANTREILPRVGRPVITYGFNEDNIR 272
QY 231 AQNIQITDKGTAFVYVYDGEYDFHFLSPQYGDHTVNLALAVIAISYLEKLDVTNIKEALE 290

DB 273 AIDVEQDGMKSHFTVLKRGREPLRLTINQGLHNVNLALAAIGVATDEGVSDDEALSRAUK 332
QY 291 TFGGVKRRFN-----ETTIANQVIVDDYAHHPREISATITETARKYKPKHVAVFPQHTF 345
DB 333 GFSVGRRFQVQGEFELGEGNVKLVDDYGHHPKEVEATIKAAQSHPDRLVMLFOPHRY 392
QY 346 SRTQAFLENEFAESLSKADRVFLCEIFGS-----IRENTGALTQ-----DLIDKIE 391
DB 393 SRTDRCDFDIEVLQVQDQLLLLEVYPAGEKPIVAGDSRTLARSIRLRGQVEPILIDPVE 452
QY 392 GASLINEDSI--NVLEQFDNAVILFMGAGDI 420
DB 453 G-----NLQNTIMQNVLQ--PNDLLITQAGNV 477

RESULT 7

US-09-540-236-2712
; Sequence 2712, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAI
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2712
; LENGTH: 493
; TYPE: PRT
; ORGANISM: M. catarrhalis
US-09-540-236-2712

Query Match 23.4%; Score 532; DB 4; Length 493;

Best Local Similarity 32.1%; Pred. No. 2.2e-43;
Matches 146; Conservative 78; Mismatches 185; Indels 46; Gaps 13;

QY 5 HFVGIGKSGMSLAQIMHDLGHEVQSGDIENYVTFEVALRNKGIKI-LPFDANNIKEDMV 63
DB 34 HFGVIGKSGMSLAQIMHDLGHEVQSGDIENYVTFEVALRNKGIKI-LPFDANNIKEDMV 92
QY 64 VIOGNAPASSHEEIVRAHQKLDVSYNDFLQIIDDQYTSVAVTGAHGTSTTGLLSHYM 123
DB 93 VVSSAIDRQNPETRAALKAHIPVRRADMLGELMYRHRHFIAGVAGHGTSTTGLLSHYM 152
QY 124 -NGDKTSFLIG-----DGTGMLPESDYFAFEACEYRHRHFLSYKPDYAIMTNIDPDHDP 177
DB 153 TEAGLDPTVYVIGKSLNASGKNAALGASRYLVAEADSDASFLSLRPMACVVTNIDEDHME 212
QY 178 -YFKDINDVDFQEMAHNVK-KGI-IAWGDDDEHLRKEADVPYIYVYFGKSDDDIYAQ 232
DB 213 TYEGSFDKLDKDTFQVFLHNLPPYGLAVGCGDDKELYAMIDDIARPVITYGLEKNDV--- 269
QY 233 NIQITDKGTAFVYVYDGEYDFHFLSPQY-----DGTVNLALAVIAISYLEKL 280
DB 270 -----QAVDVIAAGT-KTHFTVLRKDKPLPITLNPIGHNVNVALGAJTMATDEGV 320
QY 281 DVTNIKEALETFGGVKKRRFNET-----TIANQVIVDDYAHHPREISATITETARKYKPH 333
DB 321 SDKAICQAVEKFAVGRRFENNNGSYPLTDGSGDVVLLIDDYGHHPTEIANTIKAAQOSYD 380
QY 334 KEVAVFOPHPSRTOAFLENEFAESLSKADRVFLCEIFGSIRENTGALTQIDLDKIEGA 393
DB 381 RRLVMMFOPHPSRTOAFLENEFAESLSKADRVFLCEIFGSIRENTGALTQIDLDKIEGA 440
QY 394 SLINEDSTNV--LEQFDNAVILFMGAGDIQKLQNA 426
DB 441 GQVEFIVLVNDKKEQITQVLTITINANDLLMTQGA 475

RESULT 8

US-09-489-039A-12280

123 QY -----MNGDKKTSFLIGDGTGMLPESDYFAFAACEYRRHFLSKYKPDYAIMTND 172
142 Db YQAQGLDPTFVNG-----GLVKSAGTHARLGCGRYLIAEADSDASPLHLQPMVAVVTNIE 197
173 QY FDHPD-YFKDINDVDFDQEMAHNVKGLIANG-----DDHLRKLIEADVPIY--YGF 223
198 Db ADHMDTHGNFDNLKETFITFLNLP-----FGRVAMCIDDVIRSIIPKVGRIYITTYGF 253
224 QY KSDDDIYAQNTQITDKGTA-----FDVVVDGEFYDHFPLSPQYGDHVTVLNALAV 271
254 Db SEDADVRIHVE--QKGAQGFETISREDMPDIDVVLNAP-----GRNALNATRA 301
272 QY TALSILEKLDVNIKEALETFGCVKREN-----ETTIANQVIVDDVAHHPRETS 321
302 Db VAVTEGIADEHILALNFQGTGRFRDFLGNFSLHVNQGEVMLVDDYGHHPTEVD 361
322 QY ATITETARKYPHKEVAVFQPHFRTKQAFNLNFAESLSKADRVFLCEIFGS----- 373
362 Db ATIKARAGWPKRLVMLFQPHRYTRDLVEDFATVNLQVDILLTLDVYAAGEAPIPGA 421
374 QY -IRENTGALTQDLIDKIEGASLINESINVLQEPDNVILFMGAGDIQKL 423
422 Db DSRSLCIRTQRGLDPTIWSVDENISSILAGVLTDNDLVVQAGNICKI 472

RESULT 11
US-09-198-452A-978
; Sequence 978, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 978
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-978

Query Match 20.7%; Score 470; DB 4; Length 812;
Best Local Similarity 28.2%; Pred. No. 5.5e-37;
Matches 131; Conservative 89; Mismatches 181; Indels 64; Gaps 16;

4 YHFGVIGKSGMSSIAQIMHDLGHEVQSGDI-ENVYTFEVALRNKGIKLP-FDANNIKED 61
10 YHFGVIGKSGMSSIAQIMHDLGHEVQSGDI-ENVYTFEVALRNKGIKLP-FDANNIKED 67
62 MVVIOGNAPASSHEEIVRAHQKLDVVSYNDFLQIIDQYTSVAVTGAGTKTSTGLSH 121
68 AVVYSSSIAPDNVYELTATQSRRLHRAELLSQLMEGYEILVSGHGTGTSSLIRA 127
122 V-MNGDKKTSFLIGDGTGMLP-----ESDYFAFAACEYRRHFLSKYKPDYAIMTNI 171
128 IFOEAQKPSYAIG-----GLAANCLNGYSSSKI FVAEADSDGSLKHYPRAWITNI 182
172 DFPH-PDYFKDINDVDFDQEMAHNVKGLIANGDDDEHLRKI--EADVPI-----YY 221
183 DNEHLNNAAGNLNDLVQIDFSRKV-----TDLNKVFYNGDCPILKGNVQGISY 232
222 GFKSDDDIYAQNTQITDKGTAFDVVVDGEFYDHFPLSPQY-----GDHVTVLNALAVIA 273
233 GY--SPECQLHIVSYNQAW-----QSHFSFTFLGOEYQDIELNLPQGNANAAAAACG 284
274 ISYLEKLDVNIKEALETFGCVKRRFNETTIANQ-VIVDDYAHHPREISATITETARKKYP 332
285 VALTFGIDINIRKALKKESGVRHRLERKNISFSLFEDYAHHPVEVAHTLSVRDVG 344
333 HKEVAVFQPHFRTKQAFNLNFAESLSKADRVFLCEIFGSIRENTGALTQDLIDKIEG 392

345 Db LRVVIAIFQPHRFSRLBECIQTFFKPAQEADEVILTDVYSAGESPRESIIISDLAEQIRK 404
393 QY ASLI-----NEDSINVLEOF--DNVILFMGAGDIQKLQNAVLD 429
405 Db SSVVHCYVPHGDIVDYLRLNRIHDCVSLGAGNIYITIGEALXD 449

RESULT 12
US-09-328-352-4942
; Sequence 4942, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4942
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4942

Query Match 18.3%; Score 417; DB 4; Length 475;
Best Local Similarity 29.2%; Pred. No. 3.6e-32;
Matches 135; Conservative 75; Mismatches 202; Indels 50; Gaps 15;

3 YHFGVIGKSGMSSIAQIMHDLGHEVQSGDIENVYTFEVALRNKGIKLP-FDANNIK-- 59
20 HLHLIGICCTFGMGLALLARDLGHKVTGSDSVYPMSTQLENAGHELMOGYDRSHLQPH 79
60 EMVVIQGNAPASSHEEIVRAHQKLDVVSYNDFLQIIDQYTSV-AVTGAHGTSTTGL 118
80 PDLVIV-GNAMKRGIDAVEYMLNEGLPYISGPPQLADHVLQGRKVLGVAGTHGKTTTMM 138
119 LSHVMN-GDKKTSFLIGDGTGMLPES-----DYFAFAACEY-----RRHFLSKYKPD 164
139 LAWLVDQAGLNFGLIG-GVPLGFSERLGGGKYFVWEADYDSAFDFKRSKFVHYHPK 197
165 YAIMTNDIDPHDYPKINDVDFDQEMAHNV--KGLIANGDDDEHLRKI--EADVPIY 219
198 TALNLSDFDHADI FDDLAAIQKHLLVITPISGRIIAPITETHIDEVLEMGCTPVI 257
220 YXGFKSDDDIYAQNTQITDKGTAFDVVVDGEFYDHFPLSPQYGDHVTVLNALAVIAISYLEK 279
258 RTSLEANEKAALSABELISIDGSHFKVLENGVIGEVKMSMTGQHSVANALATTAQAQHV 317
280 LDVTNIKEALETFGCVKRRFNETTIANQV-IVDDYAHHPREISATITETARKKYPHKEVVA 338
318 VSLERACEALSNGFGVKKRMELLLTINGIEVYDDFAHPTAIDTLDGARKLGERRLWA 377
339 VFQPHFYS-RQAFLNEPAESLSKADRVFLCE-----IFGSIRENTGALTQIDL 386
378 IIEPRNTRMGSHKQGLAHSARLADVIWYQPEGLDQLQPVIEAATNHAQVRSRLDEI 437
387 IDKIEGASLINESINVLQEPDNVILFMG--AGDIQKLONA 426
438 IDRI-----VNEAGEGDAVIMSGNGPGGLHQKLMSA 469

RESULT 13
US-09-540-236-2891
; Sequence 2891, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAF
; FILE REFERENCE: 2709-2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 14:40:41 ; Search time 48 Seconds
(without alignments)
2561.354 Million cell updates/sec

Title: US-10-712-713-2

Perfect score: 2275

Sequence: 1 MTHYFVGIKSGMSLAQI.....GDIQLQNAYLDKLGKKNAP 437

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA.*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap.*
 - 2: /cgn2_6/ptodata/1/pubpaa/ECT_NEW_PUB.pap.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pap.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pap.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pap.*
 - 6: /cgn2_6/ptodata/1/pubpaa/ECTUS_PUBCOMB.pap.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pap.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pap.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pap.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pap.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pap.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pap.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pap.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pap.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pap.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pap.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2275	100.0	437	9	US-09-925-637-2
2	2275	100.0	437	14	US-10-084-205-2
3	2275	100.0	437	16	US-10-712-713-2
4	2274	100.0	437	12	US-10-282-122A-43780
5	2267	99.6	437	9	US-09-815-242-5297
6	2267	99.6	444	9	US-09-815-242-12293
7	1985	87.3	437	12	US-10-282-122A-71178
8	1470	64.6	436	12	US-10-282-122A-46331
9	1438.5	63.2	445	12	US-10-282-122A-57306
10	1433.5	63.0	456	9	US-09-815-242-10870
11	1395	61.4	444	12	US-10-282-122A-57651
12	1342	59.0	447	12	US-10-282-122A-60988
13	1139	50.1	442	12	US-10-282-122A-74395
14	1132.5	49.8	443	12	US-10-282-122A-72179
15	1119	49.2	444	9	US-09-815-242-13507

16	1119	49.2	444	12	US-10-282-122A-74073
17	1064	46.8	246	12	US-10-282-122A-71315
18	1043	45.8	422	9	US-09-765-272-116
19	690.5	30.4	461	12	US-10-282-122A-52628
20	641.5	28.2	458	12	US-10-282-122A-51882
21	585.5	25.7	450	12	US-10-282-122A-53624
22	574	25.2	482	12	US-10-282-122A-44665
23	559	24.6	469	12	US-10-282-122A-61427
24	548	24.1	506	12	US-10-282-122A-66060
25	545	24.0	469	12	US-10-282-122A-65124
26	538	23.6	468	12	US-10-282-122A-47281
27	532	23.4	486	12	US-10-282-122A-63073
28	509	22.4	465	12	US-10-282-122A-50582
29	507	22.3	491	12	US-10-282-122A-55678
30	506.5	22.3	468	12	US-10-282-122A-51373
31	503.5	22.1	485	9	US-09-738-626-5862
32	502	22.1	479	12	US-10-282-122A-69707
33	499	21.9	491	12	US-10-282-122A-59579
34	496	21.8	833	12	US-10-282-122A-55085
35	493.5	21.7	477	12	US-10-282-122A-54019
36	491	21.6	491	9	US-09-741-669-455
37	491	21.6	491	9	US-09-815-242-10032
38	491	21.6	491	12	US-10-282-122A-56416
39	490.5	21.6	491	12	US-10-282-122A-76022
40	490	21.5	480	9	US-09-815-242-12008
41	490	21.5	480	12	US-10-282-122A-66650
42	488.5	21.5	482	12	US-10-282-122A-68321
43	488	21.5	476	12	US-10-282-122A-73039
44	487	21.4	481	12	US-10-282-122A-76478
45	484.5	21.3	491	9	US-09-815-242-14079

ALIGNMENTS

RESULT 1

US-09-925-637-2
; Sequence 2, Application US/09925637
; Patent No. US20020103338A1
; GENERAL INFORMATION:
; APPLICANT: Choi
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
; FILE REFERENCE: PB560
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/151,933
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 08/781,986
; PRIOR FILING DATE: 1997-01-03
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-06
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patentn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-637-2

Query Match 100.0%; Score 2275; DB 9; Length 437;

Best Local Similarity 100.0%; Pred. No. 2.9e-198;

Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTHYFVGIKSGMSLAQIMHDLGHEVQSGSDIENTVFTEVALRNKGIKILPDPANNIKE 60

DB 1 MTHYFVGIKSGMSLAQIMHDLGHEVQSGSDIENTVFTEVALRNKGIKILPDPANNIKE 60

QY 61 DMVVLQGNAFASSHREIVRAHOLKLDVSVNDFLGOIIDQYTSVAVTGAKGTSTGLLS 120

Db 61 DMVVIQGNFASHEEIVRAHQKLDVVSNDFLGQIIDQYTSVAVTGAGKSTTTGLLS 120
QY 121 HVMNGDKTSFLIGDGTGMLPESDYFAFEACEYRRHFLSYKPDYALMTNIDFDPDYPFK 180
Db 121 HVMNGDKTSFLIGDGTGMLPESDYFAFEACEYRRHFLSYKPDYALMTNIDFDPDYPFK 180
QY 181 DINDVFDAFQEMAHNVKGGIIANGDDEHLRKLKIBADVPIYYGFKSDSDIYAQNIQITDKG 240
Db 181 DINDVFDAFQEMAHNVKGGIIANGDDEHLRKLKIBADVPIYYGFKSDSDIYAQNIQITDKG 240
QY 241 TAPDVVVDGEFYDHFSLSPQYGDHVTNLALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
Db 241 TAPDVVVDGEFYDHFSLSPQYGDHVTNLALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
QY 301 ETTIANQVIVDDYAHHPREISATITARKKYPHKEVAVFQPHTFSTRTQAFINEFAESLS 360
Db 301 ETTIANQVIVDDYAHHPREISATITARKKYPHKEVAVFQPHTFSTRTQAFINEFAESLS 360
QY 361 KADRVFLCEIFGSIKRENTGALTIOQLIDKIEGASLINESSINVLEQDFNAVILFMGAGDI 420
Db 361 KADRVFLCEIFGSIKRENTGALTIOQLIDKIEGASLINESSINVLEQDFNAVILFMGAGDI 420
QY 421 OKLQNAVYLDKLGKNAF 437
Db 421 OKLQNAVYLDKLGKNAF 437
RESULT 2
US-10-084-205-2
; Sequence 2, Application US/10084205
; Publication No. US20030049648A1
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil
; TITLE OF INVENTION: 37 Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB515P1
; CURRENT APPLICATION NUMBER: US/10/084,205
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/151,933
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patent In Ver. 3.1
; SEQ ID NO 2
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-084-205-2
Query Match 100.0%; Score 2275; DB 14; Length 437;
Best Local Similarity 100.0%; Pred. No. 2.9e-198;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTHYFVGIKSGMSSLAQIMHDLGHEVQSGSDIENYVTFEVALRNKGIKILPFDANNIKE 60
Db 1 MTHYFVGIKSGMSSLAQIMHDLGHEVQSGSDIENYVTFEVALRNKGIKILPFDANNIKE 60
QY 61 DMVVIQGNFASHEEIVRAHQKLDVVSNDFLGQIIDQYTSVAVTGAGKSTTTGLLS 120
Db 61 DMVVIQGNFASHEEIVRAHQKLDVVSNDFLGQIIDQYTSVAVTGAGKSTTTGLLS 120
QY 121 HVMNGDKTSFLIGDGTGMLPESDYFAFEACEYRRHFLSYKPDYALMTNIDFDPDYPFK 180
Db 121 HVMNGDKTSFLIGDGTGMLPESDYFAFEACEYRRHFLSYKPDYALMTNIDFDPDYPFK 180
QY 181 DINDVFDAFQEMAHNVKGGIIANGDDEHLRKLKIBADVPIYYGFKSDSDIYAQNIQITDKG 240
Db 181 DINDVFDAFQEMAHNVKGGIIANGDDEHLRKLKIBADVPIYYGFKSDSDIYAQNIQITDKG 240
QY 241 TAPDVVVDGEFYDHFSLSPQYGDHVTNLALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
Db 241 TAPDVVVDGEFYDHFSLSPQYGDHVTNLALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
QY 301 ETTIANQVIVDDYAHHPREISATITARKKYPHKEVAVFQPHTFSTRTQAFINEFAESLS 360
Db 301 ETTIANQVIVDDYAHHPREISATITARKKYPHKEVAVFQPHTFSTRTQAFINEFAESLS 360
QY 361 KADRVFLCEIFGSIKRENTGALTIOQLIDKIEGASLINESSINVLEQDFNAVILFMGAGDI 420
Db 361 KADRVFLCEIFGSIKRENTGALTIOQLIDKIEGASLINESSINVLEQDFNAVILFMGAGDI 420
QY 421 OKLQNAVYLDKLGKNAF 437
Db 421 OKLQNAVYLDKLGKNAF 437

RESULT 2

US-10-084-205-2
; Sequence 2, Application US/10084205
; Publication No. US20030049648A1
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil
; TITLE OF INVENTION: 37 Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB515P1
; CURRENT APPLICATION NUMBER: US/10/084,205
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/151,933
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patent In Ver. 3.1
; SEQ ID NO 2
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-084-205-2

Query Match 100.0%; Score 2275; DB 14; Length 437;
Best Local Similarity 100.0%; Pred. No. 2.9e-198;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTHYFVGIKSGMSSLAQIMHDLGHEVQSGSDIENYVTFEVALRNKGIKILPFDANNIKE 60
Db 1 MTHYFVGIKSGMSSLAQIMHDLGHEVQSGSDIENYVTFEVALRNKGIKILPFDANNIKE 60
QY 61 DMVVIQGNFASHEEIVRAHQKLDVVSNDFLGQIIDQYTSVAVTGAGKSTTTGLLS 120
Db 61 DMVVIQGNFASHEEIVRAHQKLDVVSNDFLGQIIDQYTSVAVTGAGKSTTTGLLS 120
QY 121 HVMNGDKTSFLIGDGTGMLPESDYFAFEACEYRRHFLSYKPDYALMTNIDFDPDYPFK 180
Db 121 HVMNGDKTSFLIGDGTGMLPESDYFAFEACEYRRHFLSYKPDYALMTNIDFDPDYPFK 180
QY 181 DINDVFDAFQEMAHNVKGGIIANGDDEHLRKLKIBADVPIYYGFKSDSDIYAQNIQITDKG 240
Db 181 DINDVFDAFQEMAHNVKGGIIANGDDEHLRKLKIBADVPIYYGFKSDSDIYAQNIQITDKG 240
QY 241 TAPDVVVDGEFYDHFSLSPQYGDHVTNLALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
Db 241 TAPDVVVDGEFYDHFSLSPQYGDHVTNLALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300

QY 301 ETTIANQVIVDDYAHHPREISATITARKKYPHKEVAVFQPHTFSTRTQAFINEFAESLS 360
Db 301 ETTIANQVIVDDYAHHPREISATITARKKYPHKEVAVFQPHTFSTRTQAFINEFAESLS 360
QY 361 KADRVFLCEIFGSIKRENTGALTIOQLIDKIEGASLINESSINVLEQDFNAVILFMGAGDI 420
Db 361 KADRVFLCEIFGSIKRENTGALTIOQLIDKIEGASLINESSINVLEQDFNAVILFMGAGDI 420
QY 421 OKLQNAVYLDKLGKNAF 437
Db 421 OKLQNAVYLDKLGKNAF 437
RESULT 3
US-10-712-713-2
; Sequence 2, Application US/10712713
; Publication No. US20040082002A1
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil
; TITLE OF INVENTION: 37 Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB515P1
; CURRENT APPLICATION NUMBER: US/10/712,713
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: US/10/084,205
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/151,933
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patent In Ver. 3.1
; SEQ ID NO 2
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-712-713-2
Query Match 100.0%; Score 2275; DB 16; Length 437;
Best Local Similarity 100.0%; Pred. No. 2.9e-198;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTHYFVGIKSGMSSLAQIMHDLGHEVQSGSDIENYVTFEVALRNKGIKILPFDANNIKE 60
Db 1 MTHYFVGIKSGMSSLAQIMHDLGHEVQSGSDIENYVTFEVALRNKGIKILPFDANNIKE 60
QY 61 DMVVIQGNFASHEEIVRAHQKLDVVSNDFLGQIIDQYTSVAVTGAGKSTTTGLLS 120
Db 61 DMVVIQGNFASHEEIVRAHQKLDVVSNDFLGQIIDQYTSVAVTGAGKSTTTGLLS 120
QY 121 HVMNGDKTSFLIGDGTGMLPESDYFAFEACEYRRHFLSYKPDYALMTNIDFDPDYPFK 180
Db 121 HVMNGDKTSFLIGDGTGMLPESDYFAFEACEYRRHFLSYKPDYALMTNIDFDPDYPFK 180
QY 181 DINDVFDAFQEMAHNVKGGIIANGDDEHLRKLKIBADVPIYYGFKSDSDIYAQNIQITDKG 240
Db 181 DINDVFDAFQEMAHNVKGGIIANGDDEHLRKLKIBADVPIYYGFKSDSDIYAQNIQITDKG 240
QY 241 TAPDVVVDGEFYDHFSLSPQYGDHVTNLALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
Db 241 TAPDVVVDGEFYDHFSLSPQYGDHVTNLALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
QY 301 ETTIANQVIVDDYAHHPREISATITARKKYPHKEVAVFQPHTFSTRTQAFINEFAESLS 360
Db 301 ETTIANQVIVDDYAHHPREISATITARKKYPHKEVAVFQPHTFSTRTQAFINEFAESLS 360
QY 361 KADRVFLCEIFGSIKRENTGALTIOQLIDKIEGASLINESSINVLEQDFNAVILFMGAGDI 420
Db 361 KADRVFLCEIFGSIKRENTGALTIOQLIDKIEGASLINESSINVLEQDFNAVILFMGAGDI 420
QY 421 OKLQNAVYLDKLGKNAF 437
Db 421 OKLQNAVYLDKLGKNAF 437

RESULT 4

US-10-282-122A-43780
; Sequence 43780, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43780
; LENGTH: 437
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-10-282-122A-43780

Query Match 100.0%; Score 2274; DB 12; Length 437;
Best Local Similarity 99.8%; Pred. No. 3.6e-198;
Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTHYFVGKSGMSSLAQIMHDLGHEVQSGDIENVYFTEVALRNKGIKILPFDANNIKE 60
DB 1 MTHYFVGKSGMSSLAQIMHDLGHEVQSGDIENVYFTEVALRNKGIKILPFDANNIKE 60
QY 61 DMVVIQGNAPASSHEEIVRAHQKLDVVSYNDFLQGIIDQYTSVAVTGAHGKSTTTGLLS 120
DB 61 DMVVIQGNAPASSHEEIVRAHQKLDVVSYNDFLQGIIDQYTSVAVTGAHGKSTTTGLLS 120
QY 121 HVMNGDKKTSFLIGDGTGMLPESDYFAPEACEYRRHFLSYKPDYAIMTNIDFDPDYPK 180
DB 121 HVMNGDKKTSFLIGDGTGMLPESDYFAPEACEYRRHFLSYKPDYAIMTNIDFDPDYPK 180
QY 181 DINDVFDAFQMAHNVKKGIIAWGDDEHLRKEADVPIYYGFKDSDDIYAQNIQITDKG 240
DB 181 DINDVFDAFQMAHNVKKGIIAWGDDEHLRKEADVPIYYGFKDSDDIYAQNIQITDKG 240
QY 241 TAFDVYDGEFYDHFSLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETEGGVKRRFN 300

Db 241 TAFDVYDGEFYDHFSLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETEGGVKRRFN 300
QY 301 ETTIANQVIVDDYAAHHPREISATITARKKYPHKEVAVVAFQPHTFSTRTQAFLEPAESLS 360
DB 301 ETTIANQVIVDDYAAHHPREISATITARKKYPHKEVAVVAFQPHTFSTRTQAFLEPAESLS 360
QY 361 KADRVFLCEIPGSIRENTGALTIOQLIDKIBGASLINEEDSNVLQFQDNNAVILFMGAGDI 420
DB 361 KADRVFLCEIPGSIRENTGALTIOQLIDKIBGASLINEEDSNVLQFQDNNAVILFMGAGDI 420
QY 421 OKLQNAVLDKLGKNAF 437
DB 421 OKLQNAVLDKLGKNAF 437

RESULT 5
US-09-815-242-5297
; Sequence 5297, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5297
; LENGTH: 437
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-815-242-5297

Query Match 99.6%; Score 2267; DB 9; Length 437;
Best Local Similarity 99.5%; Pred. No. 1.6e-197;
Matches 435; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTHYFVGKSGMSSLAQIMHDLGHEVQSGDIENVYFTEVALRNKGIKILPFDANNIKE 60
DB 1 MTHYFVGKSGMSSLAQIMHDLGHEVQSGDIENVYFTEVALRNKGIKILPFDANNIKE 60
QY 61 DMVVIQGNAPASSHEEIVRAHQKLDVVSYNDFLQGIIDQYTSVAVTGAHGKSTTTGLLS 120
DB 61 DMVVIQGNAPASSHEEIVRAHQKLDVVSYNDFLQGIIDQYTSVAVTGAHGKSTTTGLLS 120
QY 121 HVMNGDKKTSFLIGDGTGMLPESDYFAPEACEYRRHFLSYKPDYAIMTNIDFDPDYPK 180
DB 121 HVMNGDKKTSFLIGDGTGMLPESDYFAPEACEYRRHFLSYKPDYAIMTNIDFDPDYPK 180
QY 181 DINDVFDAFQMAHNVKKGIIAWGDDEHLRKEADVPIYYGFKDSDDIYAQNIQITDKG 240

```
Db 181 DINDVDAFQEMAHNVKGGIIAWGDDHRLKIEADVPIIYYGFKSDSDIYAQNIQITDKG 240
QY 241 TAFDVVDGFEYDHFSLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
Db 241 TAFDVVDGFEYDHFSLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
QY 301 ETTIANQVIVDDYAHHPREISATITETARKKYPHKEVVAVFQPHFTFSRTQAFINEFAESLS 360
Db 301 ETTIANQVIVDDYAHHPREISATITETARKKYPHKEVVAVFQPHFTFSRTQAFINEFAESLS 360
QY 361 KADRVFLCEIFGSIRENTGALTQDLIDKIEGASLINESINVLEQFDNAVILFMGAGDI 420
Db 361 KADRVFLCEIFGSIRENTGALTQDLIDKIEGASLINESINVLEQFDNAVILFMGAGDI 420
QY 421 QKLNAYLDKLGKNAF 437
Db 421 QKLNAYLDKLGKNAF 437

RESULT 6
US-09-815-242-12293
; Sequence 12293, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12293
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Staphylococcus aureus

US-09-815-242-12293

Query Match 99.6%; Score 2267; DB 9; Length 444;
Best Local Similarity 99.5%; Pred. No. 1.6e-197;
Matches 435; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTHYHFVGIKSGMSSLAQIMHDIQHEVQSGSDIENYVTEVALRNKGKILPFDANNIKE 60
Db 8 MTHYHFVGIKSGMSSLAQIMHDIQHEVQSGSDIENYVTEVALRNKGKILPFDANNIKE 67
QY 61 DMVVIQGNAFASSHEEIVRAHQKLDVVSYNDFLGQIIDQYTSVAVTGAHGTSTTGLLS 120
Db 68 DMVVIQGNAFASSHEEIVRAHQKLDVVSYNDFLGQIIDQYTSVAVTGAHGTSTTGLLS 127
QY 121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSKPDYAIMTNIDFDPDYFK 180
```

```
Db 128 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSKPDYAIMTNIDFDPDYFK 187
QY 181 DINDVDAFQEMAHNVKGGIIAWGDDHRLKIEADVPIIYYGFKSDSDIYAQNIQITDKG 240
Db 181 DINDVDAFQEMAHNVKGGIIAWGDDHRLKIEADVPIIYYGFKSDSDIYAQNIQITDKG 247
QY 241 TAFDVVDGFEYDHFSLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
Db 241 TAFDVVDGFEYDHFSLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 307
QY 301 ETTIANQVIVDDYAHHPREISATITETARKKYPHKEVVAVFQPHFTFSRTQAFINEFAESLS 360
Db 301 ETTIANQVIVDDYAHHPREISATITETARKKYPHKEVVAVFQPHFTFSRTQAFINEFAESLS 367
QY 361 KADRVFLCEIFGSIRENTGALTQDLIDKIEGASLINESINVLEQFDNAVILFMGAGDI 420
Db 361 KADRVFLCEIFGSIRENTGALTQDLIDKIEGASLINESINVLEQFDNAVILFMGAGDI 427
QY 421 QKLNAYLDKLGKNAF 437
Db 421 QKLNAYLDKLGKNAF 444

RESULT 7
US-10-282-122A-71178
; Sequence 71178, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71178
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis

US-10-282-122A-71178
```

Query Match	87.3%;	Score 1985;	DB 12;	Length 437;
Best Local Similarity	84.7%;	Pred. No. 7.3e-172;		
Matches 370;	Conservative 43;	Mismatches 24;	Indels 0;	Gaps 0;
QY	1	MTYHFVGLKGSWSSLAQIMHDLGHEVQGSIDIENVYFTEVALRNKGKILPPDANNIKE	60	
DB	1	MTYHFVGLKGSWSSLAQIMHDLGHEVQGSIDIENVYFTEVALRNKGKILPPDANNITK	60	
QY	61	DMVVIQGNAPFASHEEIVRAHQKLDVSVSYNDLFGQIIDQYTSVAVTGAHGKSTTGGLLS	120	
DB	61	EMVVIQGNAPFNDHEEIVRAHELKLDIIKYHDFLGHVINOYTSVAVTGAHGKSTTGGLLS	120	
QY	121	HVMNGDKTSFLIGDGTGMGLPSDSYPAFAECYRRHFLSYKPDYAIWNTIDFDPDYPFK	180	
DB	121	HVMNGDKTSFLIGDGTGMGLPSDSYPAFAECYRRHFLSYHPDYAIWNTIDFDPDYPFK	180	
QY	181	DINDVFDAFQEMAHNVKGGIIANGDDHILRKIEADVPYIYYGFKDSODDIYAQNIQITDKG	240	
DB	181	NIDDDVDAFQHMALNVKGGIIANGDDHILRKLDVDIPVYIYGFKETDDIYAKNIQITEKG	240	
QY	241	TAFDVPVVDGEFYDHFSLSPQXGDHIVLNALAVIAISYLEKLDVNTIKEALTFFGVGKREIN	300	
DB	241	TQFDVYIKGEFYDQFLSPQXGNHILNMAVIAISYLEDMNENIKALITFFGVGKREIN	300	
QY	301	ETTIANQVIVDDYAHHPREISATETARKKYPHKEVVAVPQPTFTSQTAFLNEFAESLS	360	
DB	301	ETKVSQVIVDDYAHHPREISATETARKKYPKQVVAVPQPTFTSQTAFLNEFAESLS	360	
QY	361	KADRVFLCEIFGSIRENTGALTQQDLIDKIEGASLINEDSINVLQEPDNAVILPMWGAGDI	420	
DB	361	KADQVFLCEIFGSIRENTGDLTIEDLINRDGSTLIDENSIDVLEKFDNAVILPMWGAGDI	420	
QY	421	QKLNAYDLKLGKMNAP	437	
DB	421	QKILKAYFEKLGKVNDF	437	

```

RESULT 8
US-10-282-122A-46331
? Sequence 46331, Application US/10282122A
? Publication No. US20040029129A1
? GENERAL INFORMATION:
? APPLICANT: Wang, Liangsu
? APPLICANT: Zamudio, Carlos
? APPLICANT: Malone, Cheryl
? APPLICANT: Haselbeck, Robert
? APPLICANT: Ohlsen, Kari
? APPLICANT: Zyskind, Judith
? APPLICANT: Wall, Daniel
? APPLICANT: Trawick, John
? APPLICANT: Carr, Grant
? APPLICANT: Yamamoto, Robert
? APPLICANT: Forsyth, R.
? APPLICANT: Xu, H.
? TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
? FILE REFERENCE: ELITRA 03A4
? CURRENT APPLICATION NUMBER: US/10/282,122A
? CURRENT FILING DATE: 2003-02-20
? PRIOR APPLICATION NUMBER: 60/191,078
? PRIOR FILING DATE: 2000-03-21
? PRIOR APPLICATION NUMBER: 60/206,848
? PRIOR FILING DATE: 2000-05-23
? PRIOR APPLICATION NUMBER: 60/207,727
? PRIOR FILING DATE: 2000-05-26
? PRIOR APPLICATION NUMBER: 60/230,335
? PRIOR FILING DATE: 2000-09-06
? PRIOR APPLICATION NUMBER: 60/230,347
? PRIOR FILING DATE: 2000-09-09
? PRIOR APPLICATION NUMBER: 60/242,578
? PRIOR FILING DATE: 2000-10-23
? PRIOR APPLICATION NUMBER: 60/253,625
? PRIOR FILING DATE: 2000-11-27
? PRIOR APPLICATION NUMBER: 60/257,931

```

```

; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46331
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; US-10-282-122A-46331

Query Match          64.6%; Score 1470; DB 12; Length 436;
Best Local Similarity 64.7%; Pred. No. 5 4e-125;
Matches 279; Conservative 59; Mismatches 93; Indels 0; Gaps 0;

Qy      1 MTHYHFVGIKSGMSSLAQIMHDLGHEVGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 MTHYHFVGIKGTGMSSLAQILHDMKHTVQGSDEYKRRFTQTALKEKRNISILPFDKSNVKE 60

Qy      61 DMVVIQGNAPASHSHEIIVZAHQIKLDVVSYNDFLQGIIDQYTSVAVTGAHGKTSITGLLS 120
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      61 GQVVIAGNAPPDTHESIIVAAKELNIPVHRHYHFLGLMLNQYTSVAVTGAHGKTSITGLLA 120

Qy      121 HVANGDKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDVAIMTNIDFDHPDYFK 180
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      121 HVMGQAHPTSYLIGDGTGHGVENSKYTFVEACEYRRHFLSYNPDVAIMTNIDFDHPDYFT 180

Qy      181 DINDVDFAFQEMAHNVKGGIIAWGDDEHLRKIEADVPYIYYGFKOSDDIYAQNIQITDKG 240
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      181 DINDVFSAFQEMALVKVGGIIAGDDDEELQKQAKVPVIFYFGEDNDFQARNIQKRTDG 240

Qy      241 TAPDVVVDGEFYDHFUSPOYGDHITVLNALAVIAISYLEKLDVTNIIKEALETGGVKKRRE 300
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      241 TIFDVFVRNTYDITFKITGYGNHSLNALAVIALCHYENVNDVEAVKHQLTTFEGVKRRFN 300

Qy      301 ETTIANQVIVDDYAHHPREISATIEATKKYPHKEVAVFQPHTFESRTQAFLEFAESLS 360
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      301 EKPMGEQVVIDDYAHHTPTINATIEAARQKHPEREIVAVFQHTFSRTBKFIDFEAESLS 360

Qy      361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      361 KADQVYLCDFGSARENKELTIEDLQKRIDGAELITDITTDVLKKKHGVLIIFMGAGDI 420

Qy      421 QKLQNAVYDKL 431
      |||:|||||:
Db      421 QKPEAAVYKEV 431
      |||:|||||:

RESULT 9
US-10-282-122A-57306
; Sequence 57306, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zvekund, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078

```



```
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 57306
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-10-282-122A-57306

Query Match          63.2%; Score 1438.5; DB 12; Length 445;
Best Local Similarity 60.9%; Pred. No. 4.1e-122;
Matches 259; Conservative 79; Mismatches 86; Indels 1; Gaps 1;

QY 4 YHFGVIGKSGMSSLAQIMHDLGHEVQGSDIENYVTFEVALRNKGIKILPFDANNIKEDMV 63
Db 9 YHFGVIGKSGMSSLALVHLQKGYNVQGSDEVEYFTQDLEKSGVPLFPNADNIDKDMI 68

QY 64 VIQGNAPASSHEEIVRAHQKLDVSVNDELQIIDDQYTSVAVTGAKGKTSTTGLLSHVM 123
Db 69 VIAGNAPDPDTHETETARAEIAGAEVIRYHDFIARFIEPTSIATVGTSGHKTSTTGLLAHVL 128

QY 124 NGDKKTSFLIGDGTGMLPESDYPFAFACERYRHRFLSYKPDYAIMTNIDFDPDYPKDN 183
Db 129 SGINPTSYLIGDGTGHEPDADFFAFACERYRHRFLAYSPPDYAIMTNIDFDPDYPKDN 188

QY 184 DVFPDQFQMAHNVKGGIIAWGDDEHLRKIRADVPYIYYGFKDSDDIYAQNIQITDKGTAF 243
Db 189 DVFSAFQTMHQVKKGIFAYGDDKYLRLQLESEVPVYIYGVSEEDDIQARNIQRTEGSSP 248

QY 244 DVYVDGEFDFHLSPOYGDHVTVALAVIAISYLEKLDVTNIKEALETFGGVKKRFRNETT 303
Db 249 DVYHKDDFVGHFVLPAPFGHHNNMAGLVIAVAYFEKLDQKVAEMLSFKGVKKRFRSEKK 308

QY 304 IANOVIVDDYAHHPREISATITETARKKYPHKEVAVFQPHFTSRTOAFLENEFAESLSKAD 363
Db 309 VSDMIIVDDYAHHPAEIKATIDGARQKYPDKEIIAVFQPHFTFRTIALMDFEAEALDLAD 368

QY 364 RVFLCEIFGSIRENTGALTIQDLIDKIE-GASLINEDSINVLEQFDNAVILFMGAGDIQK 422
Db 369 EVFLCNIFGSARETQGEVRIEDLGEKIQKGQVITEDNVSPLLDFENAVVFMGAGDVQK 428

QY 423 LQNAV 427
Db 429 FEQAY 433

RESULT 10
US-09-815-242-10870
; Sequence 10870, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
```

```
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10870
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10870

Query Match          63.0%; Score 1433.5; DB 9; Length 456;
Best Local Similarity 60.7%; Pred. No. 1.2e-121;
Matches 258; Conservative 79; Mismatches 87; Indels 1; Gaps 1;

QY 4 YHFGVIGKSGMSSLAQIMHDLGHEVQGSDIENYVTFEVALRNKGIKILPFDANNIKEDMV 63
Db 20 YHFGVIGKSGMSSLALVHLQKGYNVQGSDEVEYFTQDLEKSGVPLFPNADNIDKDMI 79

QY 64 VIQGNAPASSHEEIVRAHQKLDVSVNDELQIIDDQYTSVAVTGAKGKTSTTGLLSHVM 123
Db 80 VIAGNAPDPDTHETETARAEIAGAEVIRYHDFIARFIEPTSIATVGTSGHKTSTTGLLAHVL 139

QY 124 NGDKKTSFLIGDGTGMLPESDYPFAFACERYRHRFLSYKPDYAIMTNIDFDPDYPKDN 183
Db 140 SGINPTSYLIGDGTGHEPDADFFAFACERYRHRFLAYSPPDYAIMTNIDFDPDYPKDN 199

QY 184 DVFPDQFQMAHNVKGGIIAWGDDEHLRKIRADVPYIYYGFKDSDDIYAQNIQITDKGTAF 243
Db 200 DVFSAFQTMHQVKKGIFAYGDDKYLRLQLESEVPVYIYGVSEEDDIQARNIQRTEGSSP 259

QY 244 DVYVDGEFDFHLSPOYGDHVTVALAVIAISYLEKLDVTNIKEALETFGGVKKRFRNETT 303
Db 260 DVYHKDDFVGHFVLPAPFGHHNNMAGLVIAVAYFEKLDQKVAEMLSFKGVKKRFRSEKK 319

QY 304 IANOVIVDDYAHHPREISATITETARKKYPHKEVAVFQPHFTSRTOAFLENEFAESLSKAD 363
Db 320 VSDMIIVDDYAHHPAEIKATIDGARQKYPDKEIIAVFQPHFTFRTIALMDFEAEALDLAD 379

QY 364 RVFLCEIFGSIRENTGALTIQDLIDKIE-GASLINEDSINVLEQFDNAVILFMGAGDIQK 422
Db 380 EVFLCNIFGSARETQGEVRIEDLGEKIQKGQVITEDNVSPLLDFENAVVFMGAGDVQK 439

QY 423 LQNAV 427
Db 440 FEQAY 444

RESULT 11
US-10-282-122A-57651
; Sequence 57651, Application US/10282122A
; Publication No. US20040029129A1
```

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 57651

LENGTH: 444

TYPE: PRF

ORGANISM: Enterococcus faecium

US-10-282-122A-57651

Query Match 61.4%; Score 1396; DB 12; Length 444;

Best Local Similarity 60.6%; Pred. No. 3e-118;

Matches 260; Conservative 74; Mismatches 93; Indels 2; Gaps 2;

QY 4 YHFGVIGKSGMSSLAQIMHDLGHEVQSGSDIENYVTEVALRNKGKILPFDANNIKEDMV 63
DB 8 YHFGVIGKSGMSSLAQIMHDLGHEVQSGSDIENYVTEVALRNKGKILPFDANNIKEDMV 67
QY 64 VIQGNAPASSHEEIVRAHQKLDVSYNDPLGQIIDQYTSVAVTGAGKTKSTTGLLSHYM 123
DB 68 LIAGNAPDSSHEEIVRAHQKLDVSYNDPLGQIIDQYTSVAVTGAGKTKSTTGLLSHYM 127
QY 124 NGDKTSFLGDTGTMGLPESDYFAFEACEYRRHFLSPDYVAIMTNDIDFHPDYKIN 183
DB 128 SGVRPTSLYIGDTGTMGLPESDYFAFEACEYRRHFLSPDYVAIMTNDIDFHPDYKIN 187
QY 184 DVDFAFQBMHNVKGIAMGDDHRLKIEADVPLYYGPKSDIYAQNTIITDKGTAF 243
DB 188 DVFTAFQTMAGQVKKALIFAYGDDAYLRKLNKAVPLYYGYVTENDDIQARNIERTTSSAF 247
QY 244 DVYVYDGFYDFHFLSPQXGDHVTWLNALAVIAISYLEKLDVMTNKEALETFGGVKKRFNETT 303
DB 248 DVYHGFDFVGHFTVPAFGKINNALGVIAVAYEKLDELKEVAEMLTFPGVKKRFSKI 307
QY 304 IANQVIVDDYAHHPRETSATITETARKYPHKEVVAVFQPFSTRTQAFINFAESLSKAD 363
DB 308 VADMTVVDDYAHHPRETSATITETARKYPHKEVVAVFQPFSTRTQAFINFAESLSKAD 367

QY 364 RVFLCEIFGSGIRENTGALTIOQLIDKI-EGASLINEDSINVLEQFDNAVILFWGAGDIQK 422
DB 368 KVLCDIFGSGAREQGNVKIEDLAGAKIKKGGEVIKENNVSPLLDYHDAVVFWMGAGDVQK 427
QY 423 LQNYLDKX 431
DB 428 FEQAY-EKL 435

RESULT 12

US-10-282-122A-60988

Sequence 60988, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John

APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert

APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 60988

LENGTH: 447

TYPE: PRF

ORGANISM: Listeria monocytogenes

US-10-282-122A-60988

Query Match

Best Local Similarity 59.0%; Score 1342; DB 12; Length 447;

Matches 255; Conservative 64; Mismatches 108; Indels 0; Gaps 0;

QY 1 MTHYFVIGKSGMSSLAQIMHDLGHEVQSGSDIENYVTEVALRNKGKILPFDANNIKE 60
DB 1 MTHYFVIGKSGMSSLAQIMHDLGHEVQSGSDIENYVTEVALRNKGKILPFDANNIKE 60
QY 61 DMVTVQGNAPASSHEEIVRAHQKLDVSYNDPLGQIIDQYTSVAVTGAGKTKSTTGLLS 120
DB 61 GLTIIAGNAPDSSHEEIVRAHQKLDVSYNDPLGQIIDQYTSVAVTGAGKTKSTTGLLS 120
QY 121 HVMNGDKTSFLGDTGTMGLPESDYFAFEACEYRRHFLSPDYVAIMTNDIDFHPDYK 180

[illegible]

RESULT 13

```

US-10-282-122A-74395
/ Sequence 74395, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELIPRA.034A

```

Query Match	50.1%;	Score 1139;	DB 12;	Length 442;
Best Local Similarity	50.7%;	Pred. No. 7.4e-95;		
Matches 218;	Conservative 82;	Mismatches 128;	Indels 2;	Gaps 2;
QY	4	YHFGIGKSGMSSIAQIMHDI	GHEVQGS	DIENYVTEVALRNKGIKILPPDANNIKEDMV 63
DB	5	YHFGIGKSGMSALMLHQMHWK	VQGS	DVKYFTQRLGEQAGITILPFS
QY	64	VIQGNFAS	SHSEETIVRAHQLKLDVVS	YNDFLGQIIDVTSVAVTGAHGKTSITGLLSHYM 123
DB	65	LIVGNARFNENK	VEYALRHQIPKRYHDF	LQDFMKSFISFAVAGAHGKTSITGLLSHVL 124
QY	124	NGDKTSLF	LIGDGTGMGLPESD	YFAFEACEYRRHFLSYKPDYAIMTINIDFDPDYFKOIN 183
DB	125	KNITDTSYL	LIGDGTGRGSANAQYVF	ESDEYERHMPVHPYSIITNIDFDPDYFTGLA 184
QY	184	DVFDAPOEMAHNVKGI	IAMGDDSHLRLKIEADVP	IIYIYFGKSDDIYIAQNIQITDKGTAF 243
DB	185	DVRNAFNDAKQVKALFVY	GEDDBELKKIEAP	PIYYGFEENDFIAYDITRTTNGSDF 244
QY	244	DVYVDGFFYDHFLSPQ	GDHTVINALAVIAISYLEK	LDVTMIKEALETFGGVKRPFNETT 303
DB	245	KVRHQGSEVI	QGFHPYPA	YKGNILNATAVIANLFVAGIDMALVA
QY	304	IANGVIVDDVAHHPREIS	ATTETARKKYPKHEVVA	VFOPHFSRTQAFINRFAEISLAD 363
DB	305	INDTIIIDDFAHHTTEI	VAITDAARQKPSKEI	VAIFQHFTRTIALLEDPA
QY	364	RVLCEITFGS	TRE-NTGALTATQD	LIDK-IEGASLINEDSINVLEQFDNAVILFMGAGDIQ 421
DB	365	SVYLAQIYGSAREV	KGEVKVEDLAAKIIKPSQ	VTVENVSPLLDHDNAVYVFMGAGDIQ 424
QY	422	KLQNAVYD	DKL 431	
DB	425	LYXSHFEEL	434	

RESULT 14

US-10-282-122A-72179
 ; Sequence 72179, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 14:38:01 ; Search time 20 Seconds
(without alignments)
2101.785 Million cell updates/sec

Title: US-10-712-713-2

Perfect score: 2275

Sequence: 1 MTHYFVGKSGMSSLAQI.....GDIQKLNAYLDKLGKNAP 437

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*

- 1: pir1.*
- 2: pir2.*
- 3: pir3.*
- 4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2274	100.0	437	2	H89958	UDP-N-acetylmuram
2	1438	63.2	432	2	C93662	UDP-N-acetylmuram
3	1368	60.1	433	2	H84055	UDP-N-acetylmuram
4	1355	59.6	447	2	A81638	UDP-N-acetyl muram
5	1342	59.0	447	2	A81275	UDP-N-acetyl muram
6	1180	51.9	457	2	H86879	hypothetical prote
7	1119	49.2	444	2	D98043	UDP-N-acetylmuram
8	1117	49.1	444	2	B95177	UDP-N-acetylmuram
9	641.5	28.2	458	2	F97296	UDP-N-acetylmuram
10	566	24.9	495	2	C71679	UDP-N-acetylmuram
11	565.5	24.9	485	2	C97741	hypothetical prote
12	548	24.1	506	2	A81777	UDP-N-acetylmuram
13	538	23.6	468	2	H70201	UDP-N-acetylmuram
14	535	23.5	469	2	B81201	UDP-N-acetylmuram
15	534	23.5	473	2	A87565	UDP-N-acetylmuram
16	521	22.9	454	2	B70418	UDP-N-acetylmuram
17	513	22.5	471	2	A82833	UDP-N-acetylmuram
18	513	22.5	471	2	B97611	UDP-N-acetylmuram
19	500.5	22.0	494	2	A82438	UDP-N-acetylmuram
20	496	21.8	833	2	B81737	UDP-N-acetylmuram
21	492.5	21.6	471	2	A83324	UDP-N-acetylmuram
22	491	21.6	491	1	CEECAM	UDP-N-acetylmuram
23	491	21.6	491	2	G90640	UDP-N-acetyl-muram
24	491	21.6	491	2	G85491	UDP-N-acetyl-muram
25	490.5	21.6	491	2	A80518	UDP-N-acetylmuram
26	490	21.5	480	2	D83094	UDP-N-acetylmuram
27	487	21.4	481	2	C71338	probable UDP-N-ace
28	470	20.7	809	2	G86603	muramate-Ala ligas
29	470	20.7	809	2	D72022	UDP-N-acetylmuram

ALIGNMENTS									
RESULT 1									
H89958									
UDP-N-Aacerylmuramate-alanine ligase [imported] - Staphylococcus aureus (strain N315)									
C;Species: Staphylococcus aureus									
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001									
C;Accession: H89958									
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc									
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;									
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramoto, K.									
Lancet 357, 1225-1240, 2001									
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.									
A;Reference number: A89758; MUID:21311952; PMID:11418146									
A;Accession: H89958									
A;Status: preliminary									
A;Molecule type: DNA									
A;Residues: 1-437 <KUR>									
A;Cross-references: GB:BA000018; PID:g13701535; PIDN:BA842829.1; GSPDB:GN00149									
A;Experimental source: strain N315									
C;Genetics:									
A;Gene: murC									
C;Superfamily: UDP-N-acetylmuramate-alanine ligase									
Query Match 100.0%; Score 2274; DB 2; Length 437;									
Best Local Similarity 99.8%; Pred No. 1.4e-148;									
Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MTHYFVGKSGMSSLAQIMHDLGHEVQSGDIENYVTEVALNKGIKILPFDANNIKE	60						
DB	1	MTHYFVGKSGMSSLAQIMHDLGHEVQSGDIENYVTEVALNKGIKILPFDANNIKE	60						
QY	61	DMVVIQGNAPASHEHIVRAHQKLDVSNDFLQIIDQYTSVAVTGAGKSTTGLLS	120						
DB	61	DMVVIQGNAPASHEHIVRAHQKLDVSNDFLQIIDQYTSVAVTGAGKSTTGLLS	120						
QY	121	HMVNGDKTSFLGDTGMLPESDYFAFCEYRRHFLSKYKDYATMTNIDFHPDYFK	180						
DB	121	HMVNGDKTSFLGDTGMLPESDYFAFCEYRRHFLSKYKDYATMTNIDFHPDYFK	180						
QY	181	DINDVFDAFOEMAHNVKKGIIANGDDEHLRKIEADVPYIYFGKSDDIYAQNIQITDKG	240						
DB	181	DINDVFDAFOEMAHNVKKGIIANGDDEHLRKIEADVPYIYFGKSDDIYAQNIQITDKG	240						
QY	241	TAPDVYDGEFYDHFSLSPQYGDHVTNLAVIAISYLEKLDVTNKEALETFGGVKRRFN	300						
DB	241	TAPDVYDGEFYDHFSLSPQYGDHVTNLAVIAISYLEKLDVTNKEALETFGGVKRRFN	300						
QY	301	ETTIANQVIVDDYAHHPREISATETARKKYPKHEVAVFPQHTFSQTQALNEFASLS	360						
DB	301	ETTIANQVIVDDYAHHPREISATETARKKYPKHEVAVFPQHTFSQTQALNEFASLS	360						
QY	361	KADRVLCETFGSIRBNTGALTIDQIDKIDEGASLINESINVLEQFDNAVILFWGAGDI	420						

AE1638
UDP-N-acetyl muramate-alanine ligases homolog murC [imported] - Listeria innocua (strain C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AE1638
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kref, J.; Kuhn, M.; Kunat, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma, A.; Schlueker, T.; Simeos, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A;Reference number: AE1077; MUID:21537279; PMID:11679669
A;Accession: AE1638
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-447 <GLA>
A;Cross-references: GB:AL592022; PIDN:CA96877.1; PID:g16414133; GSPDB:GN00178
A;Experimental source: strain Clp11262
C;Genetics:
A;Gene: murC
C;Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match 59.6%; Score 1355; DB 2; Length 447;
Best Local Similarity 60.2%; Pred. No. 1.8e-85;
Matches 257; Conservative 63; Mismatches 107; Indels 0; Gaps 0;

QY 1 MTHYFVGKSGMSLAQIMHDLGHEVQSGSDIENYVTFEVALRNKGKILPFDANNIKE 60
DB 1 MTIYFVGKSGMSALAIHLHDKGFQVQSGDVVDKYFTQKALEEKQIPIMTFSADNIKE 60
QY 61 DMVVIQGNAPFASHEEIVRAHQKLDVVSNDFLGQIIDDQYTSVAVTGAHGKSTTGLLS 120
DB 61 GLTIAGNAPFDTHEEIERALELGLSVIRYHKFLGQLIDGYTSIAITGSHGKSTTGLLS 120
QY 121 HVNMGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYALMTNIDFHDYFK 180
DB 121 HVVGAIRPTSYLIGDGTGSGTKDAKYFALEACEYQRHFLAYKPTVATMTNIDWDHDPYFK 180
QY 181 DINDVDFAPQEMAHNVKGIILAWGDDHLKIEADVPYIYVYFGKSDSDIYAQNIQITDKG 240
DB 181 SVDDVFNAFETLGKQVKAVFALGDDVLRKSLDVPYIYFGSGSENEFQAKNVKKTG 240
QY 241 TAFDVYDGEFDFHFLSPQYGDHVTNLALAVIAISYLEKLDVTNIKEALETFGSKVRKRN 300
DB 241 TKFDVYHRDEFSLSPFEPAYGDHVNLSVIALCDYEGLPVEDVKNELATFEGVKRRFS 300
QY 301 ETTIANQVIVDDYAHHPREISATITARKKYPKKEWAVFQPHTSRTOAFLEFAESLS 360
DB 301 ITEKGNQVLVDYAHHPSEIRATVNAARQKYPKNKIVAVFQPHTRTFTFLOGFADSLN 360
QY 361 KADRVFLCEIFSGSIRENTGALTQDLIDKIEGASLINESINVLQFNDNAVILFMAGDI 420
DB 361 LADEVYLCDFGSAKREKTNLTADLAHKTKGNHIIKEHEBELLYQPEAVILFMAGDV 420
QY 421 QKLQAY 427
DB 421 QKQAA 427

RESULT 5
AE1275
UDP-N-acetyl muramate-alanine ligases homolog murC [imported] - Listeria monocytogenes (C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AE1275
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kref, J.; Kuhn, M.; Kunat, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma, A.; Schlueker, T.; Simeos, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.

A;Reference number: AE1077; MUID:21537279; PMID:11679669

A;Accession: AE1275

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-447 <GLA>

A;Cross-references: GB:NC_003210; PIDN:CA99683.1; PID:g16411034; GSPDB:GN00177

A;Experimental source: strain EGD-e

C;Genetics:

A;Gene: murC

C;Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match 59.0%; Score 1342; DB 2; Length 447;

Best Local Similarity 59.7%; Pred. No. 1.4e-84;

Matches 255; Conservative 64; Mismatches 108; Indels 0; Gaps 0;

QY 1 MTHYFVGKSGMSLAQIMHDLGHEVQSGSDIENYVTFEVALRNKGKILPFDANNIKE 60

DB 1 MTIYFVGKSGMSALAIHLHDKGFQVQSGDVVDKYFTQKALEEKQIPIMTFSADNIQE 60

QY 61 DMVVIQGNAPFASHEEIVRAHQKLDVVSNDFLGQIIDDQYTSVAVTGAHGKSTTGLLS 120

DB 61 GLTIAGNAPFDTHEEIERALELGLSVIRYHKFLGQLIDGYTSIAITGSHGKSTTGLLS 120

QY 121 HVNMGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYALMTNIDFHDYFK 180

DB 121 HVVGAIRPTSYLIGDGTGSGTKDAEYFALEACEYQRHFLAYKPTVATMTNIDWDHDPYFK 180

QY 181 DINDVDFAPQEMAHNVKGIILAWGDDHLKIEADVPYIYVYFGKSDSDIYAQNIQITDKG 240

DB 181 SVDDVFNAFETLGKQVKAVFALGDDVLRKSLDVPYIYFGSGSENEFQAKNVKKTG 240

QY 241 TAFDVYDGEFDFHFLSPQYGDHVTNLALAVIAISYLEKLDVTNIKEALETFGSKVRKRN 300

DB 241 TKFDVYHRDEFSLSPFEPAYGDHVNLSVIALCDYEGLPVEDVKNELATFEGVKRRFS 300

QY 301 ETTIANQVIVDDYAHHPREISATITARKKYPKKEWAVFQPHTSRTOAFLEFAESLS 360

DB 301 ITEKGNQVLVDYAHHPSEIRATVNAARQKYPKKEWAVFQPHTRTFTFLOGFADSLN 360

QY 361 KADRVFLCEIFSGSIRENTGALTQDLIDKIEGASLINESINVLQFNDNAVILFMAGDI 420

DB 361 LADEVYLCDFGSAKREKTNLTADLAHKTKGNHIIKEHEBELLYQPEAVILFMAGDV 420

QY 421 QKLQAY 427

DB 421 QKQAA 427

RESULT 6

H86879

hypothetical protein murC [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C;Species: Lactococcus lactis subsp. lactis

C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C;Accession: H86879

R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001

A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp

A;Reference number: A86625; MUID:21235186; PMID:11337471

A;Accession: H86879

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-457 <STO>

A;Cross-references: GB:AE005176; PID:g12725089; PIDN:AAK06138.1; GSPDB:GN00146

A;Experimental source: strain IL1403

C;Genetics:

A;Gene: murC

C;Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match 51.9%; Score 1180; DB 2; Length 457;

Best Local Similarity 52.7%; Pred. No. 1.9e-73;

Matches 227; Conservative 79; Mismatches 121; Indels 4; Gaps 4;

QY 4 YHFVGKSGMSLAQIMHDLGHEVQSGSDIENYVTFEVALRNKGKILPFDANNIKEDMV 63

Db 19 YHFTGKSGMSALMLHQMKGKQSGSDTDYFFTGQLEQADVPLLPDEKNKPEFE 78
QY 64 VIQGNAPASSHE-EIVRAHQKLDVVSNDFLGQIIDQYTSVAVTGAHGKTSITGLLSHV 122
Db 79 LIAGNAFRDNNVEIAFAHKGFPFKRHEHFLGHFMEFDSIGVAGAHKTSITGLMAHV 138
QY 123 MNGDKKTSFLIGDGTGMLPESDYFAFAEACEYRRHFLSYKPDYAIMNTNIDFDHPDYFKDI 182
Db 139 MSNIYDTSVLIGDGTGRLGESYFVFESDEYERHFMYPHYEYTIMNTNIDFDHPDYFEG 198
QY 183 NDVDFAFQMAHNVKGIILAWGDEHLRKEADVPYIYYGFKSDSDIYAQNIQITDKGTA 242
Db 199 EDVTSAFQDYANNIKKIFAYGEDVNLRLKATAPAPIYYGFEANDDYRAENLVRSTRGSS 258
QY 243 FDVYDGEFYDFHLSPOYGDHVTNLALAVIATSYLEKLDVNTNIKEALETFGGVKRPFNET 302
Db 259 FDAYFRGEKIGHFVVPAYGKHNVNALSVVAVCHNLGDMTEVADHLLTFRGKRRPFTEK 318
QY 303 TTIANQVIVDDYAHHPREISATITETARKKYPKKEVAVFQPHTEFSRTQAFNEFAESLSKA 362
Db 319 KVGTEVLIIDFAHPHTEIATLDAARQKYPDREIVAVFQPHTEFSRTQAFNEFAESLSKA 378
QY 363 DRVFLCEIFGSIRE-NTGALTIQDLIDKI-EGASLINESINVLQFDNNAVILFMWAGDI 420
Db 379 DTVYLAQIYGSAREVDHHEITAQDLADKVRKPAKVIDLONVSPLLDHDGTVYVFMWAGNI 438
QY 421 QKLNAYLDKL 431
Db 439 QKYLAF-EKL 448

RESULT 7

D98043
UDP-N-acetylmuramate-alanine ligase (EC 6.3.2.8) [imported] - Streptococcus pneumoniae
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 03-Jun-2002
C:Accession: D98043
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; H
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
J. P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: D98043
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-444 <KUR>
A:Cross-references: GB:AE007317; PIDN:AAL00177.1; PID:g15459022; GSPDB:GN00174
C:Genetics:
A:Gene: murC
C:Superfamily: UDP-N-acetylmuramate-alanine ligase
C:Keywords: ligase

Query Match 49.2%; Score 1119; DB 2; Length 444;
Best Local Similarity 49.8%; Pred. No. 2.8e-69;
Matches 213; Conservative 84; Mismatches 127; Indels 4; Gaps 4;
QY 4 YHFTGKSGMSALMLHQMKGKQSGSDYFFTGQLEQADVPLLPDEKNKPEFE 63
Db 5 YHFTGKSGMSALMLHQMKGKQSGSDYFFTGQLEQADVPLLPDEKNKPEFE 64
QY 64 VIQGNAPASSHE-EIVRAHQKLDVVSNDFLGQIIDQYTSVAVTGAHGKTSITGLLSHV 122
Db 65 LIAGNAFRDNNVEIAFAHKGFPFKRHEHFLGHFMEFDSIGVAGAHKTSITGLMAHV 124
QY 123 MNGDKKTSFLIGDGTGMLPESDYFAFAEACEYRRHFLSYKPDYAIMNTNIDFDHPDYFKDI 182
Db 125 LSHITDTSFLIGDGTGRLGESYFVFESDEYERHFMYPHYEYTIMNTNIDFDHPDYFSL 184
QY 183 NDVDFAFQMAHNVKGIILAWGDEHLRKEADVPYIYYGFKSDSDIYAQNIQITDKGT 241

Db 185 EDVFNAPNDYAKQITKGLFVYGEDAEIRKITSDAPIYYGFEASGDNFVASDILLSTTGS 244
QY 242 AFVYDGEFYDFHLSPOYGDHVTNLALAVIATSYLEKLDVNTNIKEALETFGGVKRPFNE 301
Db 245 TFTVFRQNLGQFHIPTFGRHNNATAVIGLLYTAGFDLNLVREHLKTFAGVXRRETE 304
QY 302 TTIANQVIVDDYAHHPREISATITETARKKYPKKEVAVFQPHTEFSRTQAFNEFAESLSK 361
Db 305 KIYNDTVIIDFAHPHTEIATLDAARQKYPKKEVAVFQPHTEFSRTQAFNEFAESLSK 364
QY 362 ADRVFLCEIFGSIRE-NTGALTIQDLIDKI-EGASLINESINVLQFDNNAVILFMWAGDI 419
Db 365 ADAVILAQIYGSAREVDHGVKVEDLANKKHKQVITVENVSPLLDHDNAVYVFMWAGDI 424
QY 420 IQKLNAY 427
Db 425 IQTVEYSF 432

RESULT 8

B95177
UDP-N-acetylmuramate-alanine ligase [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C:Accession: B95177
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: B95177
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-444 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK75611.1; PID:g14973011; GSPDB:GN00164; TIGR:SP4
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SPI521
C:Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match 49.1%; Score 1117; DB 2; Length 444;
Best Local Similarity 49.8%; Pred. No. 3.9e-69;
Matches 213; Conservative 84; Mismatches 127; Indels 4; Gaps 4;
QY 4 YHFTGKSGMSALMLHQMKGKQSGSDYFFTGQLEQADVPLLPDEKNKPEFE 63
Db 5 YHFTGKSGMSALMLHQMKGKQSGSDYFFTGQLEQADVPLLPDEKNKPEFE 64
QY 64 VIQGNAPASSHE-EIVRAHQKLDVVSNDFLGQIIDQYTSVAVTGAHGKTSITGLLSHV 122
Db 65 LIAGNAFRDNNVEIAFAHKGFPFKRHEHFLGHFMEFDSIGVAGAHKTSITGLLSHV 124
QY 123 MNGDKKTSFLIGDGTGMLPESDYFAFAEACEYRRHFLSYKPDYAIMNTNIDFDHPDYFKDI 182
Db 125 LSHITDTSFLIGDGTGRLGESYFVFESDEYERHFMYPHYEYTIMNTNIDFDHPDYFSL 184
QY 183 NDVDFAFQMAHNVKGIILAWGDEHLRKEADVPYIYYGFKSDSDIYAQNIQITDKGT 241
Db 185 EDVFNAPNDYAKQITKGLFVYGEDAEIRKITSDAPIYYGFEASGDNFVASDILLSTTGS 244
QY 242 AFVYDGEFYDFHLSPOYGDHVTNLALAVIATSYLEKLDVNTNIKEALETFGGVKRPFNE 301
Db 245 TFTVFRQNLGQFHIPTFGRHNNATAVIGLLYTAGFDLNLVREHLKTFAGVXRRETE 304
QY 302 TTIANQVIVDDYAHHPREISATITETARKKYPKKEVAVFQPHTEFSRTQAFNEFAESLSK 361
Db 305 KIYNDTVIIDFAHPHTEIATLDAARQKYPKKEVAVFQPHTEFSRTQAFNEFAESLSK 364
QY 362 ADRVFLCEIFGSIRE-NTGALTIQDLIDKI-EGASLINESINVLQFDNNAVILFMWAGDI 419
Db 365 ADAVILAQIYGSAREVDHGVKVEDLANKKHKQVITVENVSPLLDHDNAVYVFMWAGDI 424

Db 17 HFIGIGVGMGSGIAELLYNLGVYKQSGDLVENV--NTRKLESYGKILFLGHAEOQNTNVS 74
Qy 63 VVIQGNAPASHEEIVRAHQKLDVSVNDFLGQIIDQYTSVAVTGAKTKSTT-----116
Db 75 YVWISSAINPKNPKEALKERKIPRIIRADMLAEMLRLKCSVAVSGSHGKTTTISLACL 134
Qy 117 ---GLLSHVNG---DKTSFLIGDGTGMGLPESDYFAPEACEYRRHFLSYKPDYAIM 168
Db 135 FEAGLCPTVINGGIINKSNAYLG-----SSNYLIAEADSDATFIHIPSTIAII 186
Qy 169 TNIDFDHPDYFKDINDVFADAFQEMAHNVKKGIIAWG-----DDEHLRKIEADV---PIY 219
Db 187 TNIDPEHLDDYRDFETLIGAFERSFTNLP---FYGFVAVCCIDHKIVKLVDDITERKIV 242
Qy 220 YYGFKDSD--IYAQNIQITDKGTAFDVYD-----GEFYDHFSLSPQGDHVLNALAV 271
Db 243 TYGI-DSEDAHIIAFNINNTIASSTFDVKISPNVLGTTIIEKITIPTTGRNIMLSLA 301
Qy 272 IASVLEKLDVTNIKEALETFGCVKRRFNETTIA---NOVIVDDYAAHPREISATITETAR 328
Db 302 IAVGIELDFGKAIKNGFNFGVKRRF--TKVAEYNNASIIIDDYAAHPPEIKATLATA- 358
Qy 329 KKYPHKE---VVAVFPHTFSRTOAPLNEFASLSKADRVLCFISRENTGALTIOQ 385
Db 359 KNIANKQNGKVAIFQPHRYSRMQLFDDFMLCFADADILYITDIYAAGENPIEGITGRS 418
Qy 386 LIDKTEGASLINEDSVNLEQDFNAV-----ILFWGAGDIQKLNAYLDKLGW 433
Db 419 LVDKITKRR--HHDKANFLAELDDAVGVIIIDNAASGDMIIIMMGAGNISFPANLEGRSS 476
Qy 434 K 434
Db 477 R 477

RESULT 12
A:1777
UDP-N-acetylmuramate-alanine ligase (EC 6.3.2.8) NMA2061 [imported] - Neisseria meningitidis
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002
C:Accession: A81777
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
R.; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
A.; Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: A81777
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-506 <PAR>
A:Cross-references: GB:AL162758; GB:AL157959; NID:g7380672; PIDN:CAB85279.1; PID:g738068
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: murC
A:Superfamily: UDP-N-acetylmuramate-alanine ligase
C:Keywords: ligase

Query Match 24.1%; Score 548; DB 2; Length 506;
Best Local Similarity 32.5%; Pred. No. 5.4e-30;
Matches 151; Conservative 79; Mismatches 193; Indels 42; Gaps 14;

Qy 1 MTHYHFVGIKSGMSLAQIMHDLGHEVQSGSDIENYVTEVALRNKGKILP-EDANNIK 59
Db 43 VTNIHFVIGVGVGSGIAEVLNRLGFKVSGSDQARNATE-HLSLGIQVPGHTAEHV 101
Qy 60 EDWVVIQGNAPASHEEIVRAHQKLDVSVNDFLGQIIDQYTSVAVTGAKTKSTTGL 119
Db 102 GADVVTSTAVKENPEVVALEQOQIPVIRPALMLAEIMRRFDGIAIATGKTTITSLT 161
Qy 120 SHVMN-GDKTKTSLIGD-----GTGMGLPESDYFAPEACEYRRHFLSYKPDYAIMTNDIF 173
Db 162 ASILGAAGLDPTFVIGGKLNAAAGTAVRLCKGEYIVAEADESDASFLHLTPIMSVVTNIDE 221

Qy 174 DHPD-YFKDINDVFADAFQEMAHNVK---KGI IAWGDDHLRKL--EADVPIYYVYGFKDS 227
Db 222 DHDMDTYGHSVEKLHQAFIDFIHMPFYGKAFICI-DSEHVRAILPKVSKPYATVGLDDTA 280
Qy 228 DIVAQNIOITDKGTAFDVYV-----DGEFYDHFSLSPQGDHVLNALAVIAISYLEKLD 281
Db 281 DIVATDIEVGAQMKTVHVQMGHEGSGFEVVLNMP--GRHNVNLALAAIGVALEVGAS 338
Qy 282 VTNIKEALETFGCVKRRF---NETTIAN---QVTVDDYAAHPREISATITETARKKYPHKE 335
Db 339 VBAIQKGLLGFEGVGRFQYKDIKLPNGGCTALLVDYGHHPVEMAATLSAARGAYPEKR 398
Qy 336 VVAVFPHTFSRTOAPLNEFASLSKADRVLCFISRENTGALTIOQLDKI-----390
Db 399 LVLAFPHPRYTRDGFEDFTKVLNTVDALVLTIVYAAAGEEPIAADSARALARALRVLGK 458
Qy 391 -----EGASLINEDSVNLEQDFNAVILFWGAGDIQKLNAYLD 429
Db 459 LEPIYCNVADLPEMLNLVQ--DGDIVLWAGSINRVPAALE 501

RESULT 13
H70201
UDP-N-acetylmuramate-alanine ligase (murC) homolog - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: H70201
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: H70201
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-468 <KLE>
A:Cross-references: GB:AB001180; GB:AB000783; NID:g268875; PIDN:AAC67166.1; PID:g268876
A:Experimental source: strain 831
C:Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match 23.6%; Score 538; DB 2; Length 468;
Best Local Similarity 31.0%; Pred. No. 2.3e-29;
Matches 148; Conservative 91; Mismatches 155; Indels 84; Gaps 19;

Qy 1 MTHYHFVGIKSGMSLAQIMHDLGHEVQSGSDIENYVTEVALRNKGKILPFDANNIKE 60
Db 8 LNNIFFVGIKSGACSLACFLNKGICVEGVDSKPYTDEILSNN--KISYD--NIYE 63
Qy 61 -----DMWVVIQGNAPASHEEIVRAHQKLDVSVNDFLGQIIDQYTSVAVTG 110
Db 64 FSLKQDLSFDLIVVYSSAYNKDGLQVLEAKELNTPILSYPEALGSLSKYISIGTAGSH 123
Qy 111 GKTSTT-----GLLSHVNGDKTKTSFLIGDGTGM-GLPESDYFAPEACEYRRHFL 159
Db 124 GKTTTFLGLVLFNKLGLNPNVIVGSSVKDF--KNSAIAIG--SNIFVETCEYKKHFL 179
Qy 160 SYKPDYAIMTNDPDPDYFKDINDVFADAFQEMAHNVKKG--IIAWGDDHLRKLIE-----213
Db 180 NFPNMLILTNVDYEHVDFDKFYNEALEEAFQIINLKNQGLIINSDNNLLKIKRQIN 239
Qy 214 -ADVPYIYVYGFKDSDIYAQNIQITDKGTAFDVYVVDGEFYDHFSLSPQYGD-----HTVL 266
Db 240 RKDISIFSYGSDLSDFQISNIAVSEYFCFS-----FLGLNLVELKTVLFHNL 289
Qy 267 N-ALAVIAIS-YLEK--LDVTNIKAL-----ETFGVKRRFNETTTANQVI-VDDYAAHP 317
Db 290 NFSALLALNLFLENGKSGIFDFEALIKRIAKNYSIGIKRRVEVVKENGVIYMDYAAHP 349
Qy 318 REISATITETARKKYPHKEVVAVFQHTFSTQAFNEFAESLSKADRVLCFISRENTGALTIOQLDKI-----390
Db 350 REIKNTLFGIKNFKYKRIILDFMPTTTRTKEFFADFVEVLAAADILILHNYLSNREN 409

C;Accession: A87565

Db 244 DIYATDIENVAQMKTFTVHVQMGHGGSEFVVLNMP--GRHNVNLAAAGVALEVGAS 301

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 14:33:06 ; Search time 18 Seconds
(without alignments)
1264.148 Million cell updates/sec

Title: US-10-712-713-2

Perfect score: 2275
Sequence: 1 MTHYFVGKSGMSLAQI.....GDIQKLNAYDKLGMKNF 437

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2275	100.0	437	1	MURC_STAAM
2	2254	99.1	437	1	MURC_STAAM
3	1985	87.3	437	1	MURC_STAAP
4	1470	64.6	436	1	MURC_BACCA
5	1470	64.6	436	1	MURC_BACCA
6	1438	63.2	432	1	MURC_BACSU
7	1381.5	60.7	438	1	MURC_OCEIH
8	1368	60.1	433	1	MURC_BACHD
9	1355	59.6	447	1	MURC_LISIN
10	1342	59.0	447	1	MURC_LISMO
11	1284.5	56.5	436	1	MURC_LACPL
12	1180	51.9	443	1	MURC_LACLA
13	1139	50.1	442	1	MURC_STRP3
14	1139	50.1	442	1	MURC_STRP3
15	1135	49.9	442	1	MURC_STRP8
16	1119	49.2	444	1	MURC_STRR6
17	1117	49.1	444	1	MURC_STRRN
18	701	30.8	460	1	MURC_THETN
19	641.5	28.2	458	1	MURC_CLOAB
20	626.5	27.5	457	1	MURC_CLOPE
21	596.5	26.2	468	1	MURC_FUSNN
22	566	24.9	495	1	MURC_RICPR
23	565.5	24.9	485	1	MURC_RICCN
24	548	24.1	469	1	MURC_NEIMA
25	538	23.6	469	1	MURC_BORBU
26	535	23.5	469	1	MURC_NEIME
27	534	23.5	473	1	MURC_CAUCR
28	521.5	22.9	488	1	MURC_SHEON
29	521	22.9	454	1	MURC_AQUAE
30	513	22.5	471	1	MURC_AGRTS
31	510.5	22.4	486	1	MURC_CORGL
32	506	22.2	475	1	MURC_CHLTE
33	502	22.1	486	1	MURC_PSESM

RESULT 1
MURC_STAAM
ID MURC_STAAM STANDARD; PRT; 437 AA.
AC Q99TC4; 494 500.5 22.0 471 1 MURC_ANASP
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE UDP-N-acetylmuramate-L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE acetylmuramoyl-L-alanine synthetase).
GN MURC OR SAV1740 OR SA1561 OR MW1683.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 196620;
RN [1] _SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus".
RL Lancet 357:1225-1240 (2001)
RN [2]
RC STRAIN=Mu50 / ATCC 700699, and N315;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA".
RL Lancet 359:1819-1827 (2002).
CC -|- FUNCTION: Cell wall formation.
CC -|- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +
CC phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC -|- PATHWAY: Peptidoglycan biosynthesis.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (probable).
CC -|- SIMILARITY: Belongs to the murCDEF family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; AP003363; BAB57902.1; -;
CC EMBL; AP003364; BAB42829.1; -;

```

DR EMBL; AP004828; BAB95548.1; --
DR PIR; H89558; H89558.
DR HAMAP; MF 00046; --; 1.
DR InterPro; IPR000713; Mur_ligase.
DR InterPro; IPR004101; Mur_ligase_C.
DR InterPro; IPR005758; MurC.
DR Pfam; PF01225; Mur_ligase; 1.
DR Pfam; PF02875; Mur_ligase_C; 1.
DR TIGRFAMs; TIGR01082; murC; 1.
DR TIGRFAMs; TIGR01082; murC; 1.
DR KEGG; ATP-binding; Cell division; Cell wall;
DR KEGG; Peptidoglycan synthesis; Complete proteome.
DR KEGG; NP_BIND 108 114 ATP (POTENTIAL).
FT NP_BIND 108 114
SQ SEQUENCE 437 AA; 49174 MW; 3349943079D87355 CRC64;

Query Match 100.0%; Score 2274; DB 1; Length 437;
Best Local Similarity 99.8%; Pred. No. 1.7e-142;
Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTHVFGVIGKSGMSSLAQIMHDLGHEVQGSDIENYVFEVALRNKGIKILPFDANNIKE 60
Db 1 MTHVFGVIGKSGMSSLAQIMHDLGHEVQGSDIENYVFEVALRNKGIKILPFDANNIKE 60
QY 61 DMVVIQGNAFASSHBEIVRAHQIKLDVVSYNDFLQGIIDQYTSVAVTGAGKSTTGLLS 120
Db 61 DMVVIQGNAFASSHBEIVRAHQIKLDVVSYNDFLQGIIDQYTSVAVTGAGKSTTGLLS 120
QY 121 HVMGDKKTSFLIGDGTGMLPESDYFAFEACEYRHRHFLSYKPDYALMTNIDFDHPDYFK 180
Db 121 HVMGDKKTSFLIGDGTGMLPESDYFAFEACEYRHRHFLSYKPDYALMTNIDFDHPDYFK 180
QY 181 DINDVDFAFQEAHNVKKGIIAWGDDEHLRKEADVPYIYVYGFKSDDIYAQNIQITDKG 240
Db 181 DINDVDFAFQEAHNVKKGIIAWGDDEHLRKEADVPYIYVYGFKSDDIYAQNIQITDKG 240
QY 241 TAFDVVDGFEYDFHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
Db 241 TAFDVVDGFEYDFHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
QY 301 ETTIANQVIVDDYAHHPREISATITARKKYPHKEVAVFQPHFTSRTQAFNEFAESLS 360
Db 301 ETTIANQVIVDDYAHHPREISATITARKKYPHKEVAVFQPHFTSRTQAFNEFAESLS 360
QY 361 KADRVFLCEIFGSIKRENTGALTQDLIDKIEGASLINESINVLEQFDNAVILFMGAGDI 420
Db 361 KADRVFLCEIFGSIKRENTGALTQDLIDKIEGASLINESINVLEQFDNAVILFMGAGDI 420
QY 421 QKLNAYLDKLGKNAF 437
Db 421 QKLNAYLDKLGKNAF 437

RESULT 2
MURC STAAU STANDARD; PRT; 437 AA.
AC Q31211;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE acetylmuramoyl-L-alanine synthetase).
DE MURC.
GN Staphylococcus aureus.
OS Bacteria; Firmicutes; Bacillales; Staphylococcus.
OC NCBI_TaxID=1280;
OX [1]
RN SEQUENCE FROM N.A.
RA Lowe A.M., Deresiewicz R.L.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Cell wall formation (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine +
CC phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC -!- PATHWAY: Peptidoglycan biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

```

```

CC -!- SIMILARITY: Belongs to the murCDEF family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/ebis-sib.ch/announcements)
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF034076; AAB87090.1; --
DR HAMAP; MF 00046; --; 1.
DR InterPro; IPR000713; Mur_ligase.
DR InterPro; IPR004101; Mur_ligase_C.
DR InterPro; IPR005758; MurC.
DR Pfam; PF01225; Mur_ligase; 1.
DR Pfam; PF02875; Mur_ligase_C; 1.
DR TIGRFAMs; TIGR01082; murC; 1.
DR TIGRFAMs; TIGR01082; murC; 1.
DR KEGG; ATP-binding; Cell division; Cell wall; Cell division; Ligase;
DR KEGG; NP_BIND 108 114 ATP (POTENTIAL).
FT NP_BIND 108 114
SQ SEQUENCE 437 AA; 49176 MW; E2FC7D79B6CF8361 CRC64;

Query Match 99.1%; Score 2255; DB 1; Length 437;
Best Local Similarity 98.9%; Pred. No. 2.9e-141;
Matches 432; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MTHVFGVIGKSGMSSLAQIMHDLGHEVQGSDIENYVFEVALRNKGIKILPFDANNIKE 60
Db 1 MTHVFGVIGKSGMSSLAQIMHDLGHEVQGSDIENYVFEVALRNKGIKILPFDANNIKE 60
QY 61 DMVVIQGNAFASSHBEIVRAHQIKLDVVSYNDFLQGIIDQYTSVAVTGAGKSTTGLLS 120
Db 61 DMVVIQGNAFASSHBEIVRAHQIKLDVVSYNDFLQGIIDQYTSVAVTGAGKSTTGLLS 120
QY 121 HVMGDKKTSFLIGDGTGMLPESDYFAFEACEYRHRHFLSYKPDYALMTNIDFDHPDYFK 180
Db 121 HVMGDKKTSFLIGDGTGMLPESDYFAFEACEYRHRHFLSYKPDYALMTNIDFDHPDYFK 180
QY 181 DINDVDFAFQEAHNVKKGIIAWGDDEHLRKEADVPYIYVYGFKSDDIYAQNIQITDKG 240
Db 181 DINDVDFAFQEAHNVKKGIIAWGDDEHLRKEADVPYIYVYGFKSDDIYAQNIQITDKG 240
QY 241 TAFDVVDGFEYDFHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
Db 241 TAFDVVDGFEYDFHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
QY 301 ETTIANQVIVDDYAHHPREISATITARKKYPHKEVAVFQPHFTSRTQAFNEFAESLS 360
Db 301 ETTIANQVIVDDYAHHPREISATITARKKYPHKEVAVFQPHFTSRTQAFNEFAESLS 360
QY 361 KADRVFLCEIFGSIKRENTGALTQDLIDKIEGASLINESINVLEQFDNAVILFMGAGDI 420
Db 361 KADRVFLCEIFGSIKRENTGALTQDLIDKIEGASLINESINVLEQFDNAVILFMGAGDI 420
QY 421 QKLNAYLDKLGKNAF 437
Db 421 QKLNAYLDKLGKNAF 437

RESULT 3
MURC STAAU STANDARD; PRT; 437 AA.
AC Q8CS47;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE acetylmuramoyl-L-alanine synthetase).
DE MURC OR SE1413.
GN Staphylococcus epidermidis.
OS Bacteria; Firmicutes; Bacillales; Staphylococcus.
OC NCBI_TaxID=1282;
OX

```



```
SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RC MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlant J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -!- FUNCTION: Cell wall formation.
CC -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +
CC phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC -!- PATHWAY: Peptidoglycan biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the murCDEF family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AL596169; CAC96877.1; --
CC PIR: AE1638; AE1638.
CC HAMAP: MF 00046; -- 1.
CC Listinist: LINO1646; --
CC InterPro: IPR000713; Mur_ligase.
CC InterPro: IPR004101; Mur_ligase_C.
CC InterPro: IPR005758; MurC.
CC Pfam: PF02225; Mur_ligase; 1.
CC TIGRfams: TIGR01082; murC; 1.
CC Ligase; ATP-binding; Cell division; Cell wall;
CC Peptidoglycan synthesis; Complete proteome.
CC NP_BIND 108 114 ATP (POTENTIAL).
CC SEQUENCE 447 AA; 50036 MW; AFA793A968338904 CRC64;

Query Match 59.6%; Score 1355; DB 1; Length 447;
Best Local Similarity 60.2%; Pred. No. 5e-82;
Matches 257; Conservative 63; Mismatches 107; Indels 0; Gaps 0;

QY 1 MTHYFVGIKSGMSLAQIMDLGHEVQSGDIENVVTFEVALRNKGILKLPDANNIKE 60
DB 1 MTIYHFVGIKSGMSALAQILHDKGFQVQSGVDKYFFTKALEEQIIPMTFSADNIKE 60
QY 61 DMWVIOGNFASASSEIIVRAHOKLDVSYNDPFLGQIIDQYTSVAVTGAHKSTSTGLLS 120
DB 61 GUTIIAGNAPDPTHEIERANEINLPVIRYHKLQGLIDGYTSIATGSHGKSTTGLLS 120
QY 121 HVMNGDKKTSFLIGDGTGMLPESDYFAFEACEYRRHFLSYKPDYAIMTNDIDHPDYFK 180
DB 121 HVVGAIKRTSYLIGDGTGSGTKDAKYFALEACEYRHHFLAYKPYTAIMTNDIDHPDYFK 180
QY 181 DINVDVDFAPQMAHNVKGIILAWGDDEHURKTEADVPYIYFGKSDSDIYAQNIQTIDKG 240
DB 181 SVDDVFNAFETLGGKQVKKAVFALGDDVELRKLSDVPYIYFGSGSENFQAKNVKKEITG 240
QY 241 TAFDYVVDGEFDHFLSPQYGHDTVLNALAVTASIVLEKLDVNTKEALETFGKVKREN 300
DB 241 TKFDYHRRDEFASFEPAYGDNHVNLSVIALCDYGLPVDVDEVKELKTFEGVKRRFS 300
QY 301 ETTIANQIVDDYAHHPREISATIEATKYPKHEVAVVQPHPTFRKTOAFLNFAESLS 360
DB 301 ITEKGNQVLVDYAHHPSEIRATVNRARQPNKPKIVAVQPHPTFRTRFFLGFPADSLN 360

361 KADRVFLCEIFGSIRENTGALTITQDLKIDKIEGASLINESINVLEQFNVAIVLFGAGDI 420
361 LADEVYLCDFGSAKREKTNLTIALHAKTKGNHIIKEEHTTELLQYPEAVILFMGAGDV 420
QY 421 OKLONAY 427
DB 421 QKQQAAY 427

RESULT 10
MURC_LISMO STANDARD; PRT; 447 AA.
ID MURC_LISMO
AC Q8Y6S8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE acetylmuramoyl-L-alanine synthetase).
GN MURC OR LMO1605.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;

[1]
SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RC MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlant J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -!- FUNCTION: Cell wall formation.
CC -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +
CC phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC -!- PATHWAY: Peptidoglycan biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the murCDEF family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AL596169; CAC96877.1; --
CC PIR: AE1638; AE1638.
CC HAMAP: MF 00046; -- 1.
CC Listinist: LINO1646; --
CC InterPro: IPR000713; Mur_ligase.
CC InterPro: IPR004101; Mur_ligase_C.
CC InterPro: IPR005758; MurC.
CC Pfam: PF02225; Mur_ligase; 1.
CC TIGRfams: TIGR01082; murC; 1.
CC Ligase; ATP-binding; Cell division; Cell wall;
CC Peptidoglycan synthesis; Complete proteome.
CC NP_BIND 108 114 ATP (POTENTIAL).
CC SEQUENCE 447 AA; 50036 MW; AFA793A968338904 CRC64;

Query Match 59.6%; Score 1355; DB 1; Length 447;
Best Local Similarity 60.2%; Pred. No. 5e-82;
Matches 257; Conservative 63; Mismatches 107; Indels 0; Gaps 0;

QY 1 MTHYFVGIKSGMSLAQIMDLGHEVQSGDIENVVTFEVALRNKGILKLPDANNIKE 60
DB 1 MTIYHFVGIKSGMSALAQILHDKGFQVQSGVDKYFFTKALEEQIIPMTFSADNIKE 60
QY 61 DMWVIOGNFASASSEIIVRAHOKLDVSYNDPFLGQIIDQYTSVAVTGAHKSTSTGLLS 120
DB 61 GUTIIAGNAPDPTHEIERANEINLPVIRYHKLQGLIDGYTSIATGSHGKSTTGLLS 120
QY 121 HVMNGDKKTSFLIGDGTGMLPESDYFAFEACEYRRHFLSYKPDYAIMTNDIDHPDYFK 180
DB 121 HVVGAIKRTSYLIGDGTGSGTKDAKYFALEACEYRHHFLAYKPYTAIMTNDIDHPDYFK 180
QY 181 DINVDVDFAPQMAHNVKGIILAWGDDEHURKTEADVPYIYFGKSDSDIYAQNIQTIDKG 240
DB 181 SVDDVFNAFETLGGKQVKKAVFALGDDVELRKLSDVPYIYFGSGSENFQAKNVKKEITG 240
QY 241 TAFDYVVDGEFDHFLSPQYGHDTVLNALAVTASIVLEKLDVNTKEALETFGKVKREN 300
DB 241 TKFDYHRRDEFASFEPAYGDNHVNLSVIALCDYGLPVDVDEVKELKTFEGVKRRFS 300
QY 301 ETTIANQIVDDYAHHPREISATIEATKYPKHEVAVVQPHPTFRKTOAFLNFAESLS 360
DB 301 ITEKGNQVLVDYAHHPSEIRATVNRARQPNKPKIVAVQPHPTFRTRFFLGFPADSLN 360
```

```
QY 1 MTHYFVGKSGMSLAQIMHDLGHEVQSGDIENYVTFEVALRNKGKIKLPDANNIKE 60
Db 1 MTHYFVGKSGMSALAIHDKGFQVQGVQVDFYFQKALEEKQIPIMTFSADNIQE 60
QY 61 DMVVIQGNFASHEIIVRAHOLKLDVSYNDPLGQIIOQYTSVAVTGAGTKSTTGLLS 120
Db 61 GLTIAGNAFPDTHIEIERALELGLSVIRYHFKPLGLQIDYTSIATGSHGKTSTTGLLS 120
QY 121 HVMGDKKTSFLIGDGTGMLPESDYFAFEACEYRHFELSYKPDYAIMTNIDPDHDPYFK 180
Db 121 HVVGAIPTSYLIGDGTGSGTDAEYFALEACEYQRHFELAYKPYAIMTNIDPDHDPYFK 180
QY 181 DINDVDFAPQMAHNVKGIILAWGDDEHRLKRIEADVPYIYVYGFKSDSDIYAQNIQITDKG 240
Db 181 SVDDVFNATETLGKQKVAVFAGDDAEALRKLTLDPILYFGGEENEFOAKNVIKETIG 240
QY 241 TAPDVVDGFEYDHFHLSPOYGDHTVLNALAVTASILEKLDVNIKEALTFEGVKRRFN 300
Db 241 TKFDVYHREELSSFEIPAYGDHNVNLALSVIALCDYEGLPVEDVKNELKTFEGVKRRFS 300
QY 301 ETTIANQVIVDDYAHHPREISATITETARKKYPHKEVAVFQPHTFSTQAFINEFAESLS 360
Db 301 ITEKQNVLDVDDYAHHPREISATITETARKKYPHKEVAVFQPHTFSTQAFINEFAESLS 360
QY 361 KADRYFLCEIFSGIRENTGALTITQDLIDKTEGASLINESINVLQPDNAVILFPMGAGDI 420
Db 361 LADEVTLCDIFGSAAREKTNLTITADLAHKTKGNHIIKEEHTTELLKYPEAVILFPMGAGDV 420
QY 421 QKLNAY 427
Db 421 QKFOAY 427

RESULT 11
MURC_LACPL
ID MURC_LACPL STANDARD; PRT; 436 AA.
AC Q88WZ5;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE UDP-N-acetylmuramate-L-alanine ligase (EC 6.3.2.8) (UDP-N-
acetylmuramoyl-L-alanine synthetase).
GN MURC OR IP_1462
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1590;
RN [1]_TaxID=1590;
RP SEQUENCE FROM N.A.
RC STRAIN=NCIMB 8826 / WCFS1;
RX MEDLINE=22480296; PubMed=12566566;
RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
RA De Vos W.M., Siezen R.J.;
RT "Complete genome sequence of Lactobacillus plantarum WCFS1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
CC -!- FUNCTION: Cell wall formation.
CC -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +
phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC -!- PATHWAY: Peptidoglycan biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the murCDF family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; AL935256; CAD63921.1; -.
DR HAMAP; MF 00046; -.
DR InterPro; IPR000713; Mur_ligase.
DR InterPro; IPR004101; Mur_ligase_C.
DR Pfam; PF01225; Mur_ligase; 1.
DR Pfam; PF02875; Mur_ligase_C; 1.
KW Ligase; ATP-binding; Cell division; Cell wall;
KW Peptidoglycan synthesis; Complete proteome.
FT NP_BIND 111 117 ATP (POTENTIAL).
SQ SEQUENCE 436 AA; 48750 MW; 25A582D850238CD4 CRC64;

Query Match 56.5%; Score 1284.5; DB 1; Length 436;
Best Local Similarity 55.7%; Pred. No. 2.1e-77;
Matches 240; Conservative 75; Mismatches 115; Indels 1; Gaps 1;

QY 2 THYHFVGKSGMSLAQIMHDLGHEVQSGDIENYVTFEVALRNKGKIKLPDANNIKE 61
Db 5 TVYFVGKSGMSLALILHDKGYQVSGSDIEQYTFQKGLAAAGIKMLPSEDNIREG 64
QY 62 MVIQGNFASHEIIVRAHOLKLDVSYNDPLGQIIOQYTSVAVTGAGTKSTTGLLS 121
Db 65 LTVIAGNSFTDDHPEIKAREMGLPVYRHYHFLGKLMGFTSIGVAGTHGKTSTTGLLS 124
QY 122 VMGDKKTSFLIGDGTGMLPESDYFAFEACEYRHFELSYKPDYAIMTNIDPDHDPYFK 181
Db 125 VLSHIAPTSYLGDTGKTPDARFFVFEADBYRRHFVAYHPDYAIMTNVDPDHDPYFK 184
QY 182 INDVDFAPQMAHNVKGIILAWGDDEHRLKRIEADVPYIYVYGFKSDSDIYAQNIQITDKG 241
Db 185 LADVQSAFQFGNQVKGKIFAWGDDESLRHLDVPTPIYVYGTNDRDDDFQAVNIKTKGS 244
QY 242 AFDVYVDGFEYDHFHLSPOYGDHTVLNALAVTASILEKLDVNIKEALTFEGVKRRFN 301
Db 245 SFEVYHDSLGKFEIPLFGEHNVNLSTAVSYFEKVNLDREILDFSGVKRRFSE 304
QY 302 TTTIANQVIVDDYAHHPREISATITETARKKYPHKEVAVFQPHTFSTQAFINEFAESLSK 361
Db 305 HQVGDVWVIDDYAHHPSEIKATLDAARQYKPKELAVFQPHTFSTKALMDGFAASLSK 364
QY 362 ADRVFLCEIFSGIRENTGALTITQDLIDKI -EGASLINESINVLQPDNAVILFPMGAGDI 420
Db 365 ADHVFLTNIFSSAREKSGDVSSKDLAKLPNGGEITTTDDMSALTAYHNAVAVFPMGAGDI 424
QY 421 QKLNAYLIDLK 431
Db 425 QKYEKIYEDQM 435

RESULT 12
MURC_LACLA
ID MURC_LACLA STANDARD; PRT; 443 AA.
AC Q9CE10;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE UDP-N-acetylmuramate-L-alanine ligase (EC 6.3.2.8) (UDP-N-
acetylmuramoyl-L-alanine synthetase).
GN MURC OR IL2040.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Manger S., Jaillon O., Malarre K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
CC -!- FUNCTION: Cell wall formation.
CC -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +
```

```

CC phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC -!- PATHWAY: Peptidoglycan biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the murCDEF family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; AE006433; AAK06138.1; ALT_INIT.
CC HAMAP; MF 00046; -; 1.
CC DR InterPro; IPR000713; Mur_ligase.
CC DR InterPro; IPR004101; Mur_ligase_C.
CC DR InterPro; IPR005758; MurC.
CC DR Pfam; PF01225; Mur_ligase_1.
CC DR Pfam; PF02875; Mur_ligase_C; 1.
CC DR TIGRFAMs; TIGR01082; murC; 1.
CC DR Ligase; ATP-binding; Cell division; Cell wall;
CC KW Peptidoglycan synthesis; Complete proteome.
CC FW NP_BIND 110 116 ATP (POTENTIAL).
CC SQ SEQUENCE 443 AA; 50050 MW; ABEDAB52E9638CA CRC64;
CC -----
CC Query Match 51.9%; Score 1180; DB 1; Length 443;
CC Best Local Similarity 52.7%; Pred. No. 1.6e-70;
CC Matches 227; Conservative 79; Mismatches 121; Indels 4; Gaps 4;
CC -----
CC QY 4 YHVGIGKSGMSLIAQIMHDLGHEVQSGSDIENYVTFEVALRNKGIKILPFDANNIKEDMV 63
CC Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 5 YHPTGIGKSGMSALALMLHQMKGKVGSGSDTDYFTQGLEQADVPLLPFDENIKPEFE 64
CC QY 64 VLOGNFASSHE-EIVRAHQCLKLDVSVYNDFLGOIIDQYTSVAVTGAGKSTTGLLSHV 122
CC Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 65 LIAGNAFRDNNVVEIAFAHKNRPPFKYHBEFLGHFMEDFTSIGVAGAHKSTTIGMLAHV 124
CC QY 123 MNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIIDFDHPDYFNDI 182
CC Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 125 MSNIVDTSYLLIGDTGRGIGSEYFVFESEYERHFFPYEPTIMTNIIDFDHPDYFEGI 184
CC QY 183 NDVDFAPQEAHNVKGGIIANGDDEHLRKIEADVPIYYVYFGKSDDIYAQNIQITDKGTA 242
CC Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 185 EDVTSAPQDYANNIKKGIYAGYEDVNLRLKTAKAPIYYVYFPEANDDYRAENLVRSTRGSS 244
CC QY 243 FDYVVGGEFYDHLSPQGDHTVLNALAVTAISVLEKLDVTNIKEALETFGGVKRRRNET 302
CC Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 245 FDAYFRGEKIGHFVVPAYGKHNVNLNLSVAVCHNLGLDTEVADHLITFGVKKRRITEK 304
CC QY 303 TIANQVITDDYAHHPRIISATITARKYPKPKYEVAVFQPHTFTSRTQAFINFEFAESLSKA 362
CC Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 305 KVGETVIIDFAHPTTEIATLDAARQKYPDREIVAVFQPHTFTRTIAFADEFAEVLDDHA 364
CC QY 363 DRVFLCBIFGSIRE-NTGALTIQDLIDKI-EGASLINESDINVLQEPDNVILPMGAGDI 420
CC Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 365 DTVVLAQIYGSAREVDHHEHTAQDLADKVRKPAKVIDLDNVSPLLDHDHGRGVYVFMGAGNI 424
CC QY 421 QKLNAYLIDKL 431
CC Db ||| ||| |||
CC 425 QKYELAF-EKL 434
CC -----
CC RESULT 13
CC MURC_STRP3 STANDARD; PRT; 442 AA.
CC AC Q8KBJ5;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC DE UDP-N-acetylmuramate-L-alanine ligase (EC 6.3.2.8) (UDP-N-
CC DE acetylmuramoyl-L-alanine synthetase).
CC GN MURC OR SPYM3 0252 OR SP51607.

```

OS	Streptococcus pyogenes (serotype M3).	
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;	
OC	Streptococcus.	
OX	NCBI_TaxID=198466;	
CC	[1]	
CC	SEQUENCE FROM N.A.	
RP	STRAIN=MGAS315 / Serotype M3;	
RX	MEDLINE=42133808; PubMed=1212206;	
RA	Bares S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,	
RA	Mammarilla N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,	
RA	Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,	
RA	Schlievert P.M., Musser J.M.;	
RT	"Genome sequence of a serotype M3 strain of group A Streptococcus:	
RT	phage-encoded toxins, the high-virulence phenotype, and clone	
RT	emergence.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).	
RN	[2]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN=SSI-1 / Serotype M3;	
RX	MEDLINE=422683278; PubMed=12799345;	
RA	Okagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,	
RA	Akahaishi N., Kawabata S., Yanazaki K., Shiba T., Yasunaga T.,	
RA	Hayashi H., Hattori M., Hamada S.;	
RT	"Genome sequence of an M3 strain of Streptococcus pyogenes reveals a	
RT	large-scale genomic rearrangement in invasive strains and new insights	
RT	into phage evolution.";	
RL	Genome Res. 13:1042-1055 (2003).	
CC	-1- FUNCTION: Cell wall formation.	
CC	-1- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +	
CC	phosphate + UDP-N-acetylmuramoyl-L-alanine.	
CC	-1- SUBCELLULAR LOCATION: Biosynthesis.	
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).	
CC	-1- SIMILARITY: Belongs to the murCDEF family.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (see http://www.isb-sib.ch/announcement/).	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; A014141; AAM78859.1; -	
DR	EMBL; AP005146; BAC64702.1; -	
DR	HMAP; MF 00046; -; 1.	
DR	InterPro; IPR000713; Mur_ligase.	
DR	InterPro; IPR004101; Mur_Ligase_C.	
DR	InterPro; IPR005758; MurC.	
DR	Pfam; PF01225; Mur_ligase; 1.	
DR	Pfam; PF02875; Mur_ligase_C; 1.	
DR	TIGRFAMs; TIGR01082; murC; 1.	
KW	Ligase; ATP-binding; Cell division; Cell wall;	
KW	Peptidoglycan synthesis; Complete proteome.	
FT	NE_BIND 109 115 ATP (POTENTIAL).	
SQ	SEQUENCE 442 AA; 43622 MW; C0216CAFEF922A03 CRC64;	
	Query Match 50.1%; Score 1139; DB 1; Length 442;	
	Beat Local Similarity 50.7%; Pred. No. 7.9e-68;	
	Matches 218; Conservative 82; Mismatches 128; Indels 2; Gaps 2;	
Qy	4 YHFGVIGKSGMSLAQIMHDLGHEVQSGDTEYVFTVALRNKGIKILPFDANNIKEDMV 63	
Db	5 YHFTGIGKSGMSALALHMQHGKHVQSGDYKYYFTQRLGEQAGITLTFPESDITPME 64	
Qy	64 VIQGNFASHEEIVRAHQKLDVVSYNDFLQGLIDQYTSVAVTGAHGKSTTGLLSHVM 123	
Db	65 LIVGNAFRNKKEVAYALRHQIPKRVYHDFLGDFGFMKSFISFVAGAGKSTTGLLSHVL 124	
Qy	124 NGDKKTSFLIGDGTGCMGLPESDYFAFACRYRRHFLSKPDYAIWNTIDFDHPDYFKDN 183	
Db	125 KNITDTSYLLDGTGGRGSANAQVFVESDYERHFMEYHPEYSIITNIDFDHPDYFTGIA 184	
Qy	184 DVFDAPQEMAHNVKKGIIAWGDDEHLRKIRADVPYIYYGFKOSDIIYAQNIIQTDKGTAF 243	

Db 185 DVNRAFNDYAKQVKKALFYVGGDDDELKKTIEAPAPIYVYGFEGNDPIAYDITRTTNGSDF 244
QY 244 DYVDGEFYDHLFSQYGDHTVNLAVLAIATSLVLEKLDVTNIKEALETFGGVKKRRNETT 303
Db 245 KYHQGEVIGQFHVPAVYKGNILNATAVIANLFAVAGIDMALVADHLKTFSGVKKRRFTEKI 304
QY 304 IANQVIVDDYAHHPREISATITETARKKYPHKEVAVFQPHTFSTRTOAFNLNEFAESLSKAD 363
Db 305 INDITIIDFAHPTETIVATIDARQKYSKEIVAFQPHTFSTRTOAFNLNEFAESLSKAD 364
QY 364 RVFLCEIFGSIRE-NTGALTIQDLIDK-IEGASLINESINVLQDFNAVILFMGAGDIQ 421
Db 365 SVYLAQIYGSAREVDKGEVREDLAAKIKPSQVTVENVSPLLDHDNAVYVFMGAGDIQ 424
QY 422 KIQNAVLDKL 431
Db 425 LYEHSEFEELL 434

RESULT 14

MURC_STRPY
ID MURC_STRPY STANDARD; PRT; 442 AA.
AC Q9A1C7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-acetylmuramoyl-L-alanine synthetase).
GN MURC OR SPV0345.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J., Quan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
BL Proc Natl Acad Sci U S A. 98:4658-4663(2001).
CC -1- FUNCTION: Cell wall formation.
CC -1- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP + phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC -1- PATHWAY: Peptidoglycan biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: Belongs to the murCDEF family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AE006499; AKK33396.1; --
DR HAMAP; MF_00046; -- 1.
DR InterPro; IPR000713; Mur_ligase.
DR InterPro; IPR004101; Mur_ligase_C.
DR InterPro; IPR005758; MurC.
DR Pfam; PF01225; Mur_ligase; 1.
DR Pfam; PF02875; Mur_ligase_C; 1.
DR TIGRFAMs; TIGR01082; murC; 1.
DR Ligase; ATP-binding; Cell division; Cell wall;
KW Peptidoglycan synthesis; Complete proteome.
FT NP_BIND 109 115 ATP (POTENTIAL).
SQ SEQUENCE 442 AA; 49594 MW; C02175B7FEB3303 CRC64;

Query Match 50.1%; Score 1139; DB 1; Length 442;

Best Local Similarity 50.7%; Pred. No. 7.9e-68;
Matches 218; Conservative 82; Mismatches 128; Indels 2; Gaps 2;
QY 4 YHFVIGKSGMSSLAQIMHDLGHEVQGSIDENYVTEVALRNGKIKILPFDANNIKEDMV 63
Db 5 YHFVIGKSGMSALALMLHQHGVQGSDEKYYFTQRGLEQAGITILPFSDDNITPDME 64
QY 64 VIQGNAPASSHEEIVRAHQKLDVVSNDYFLGOIQDYTSVAVTGAHKSTSTGLLSHVM 123
Db 65 LIVGNAFRENNKAVAYALRHQIPFKRYHDFLGDFFMKSFISFAVAGAHGKTSTGLLSHVL 124
QY 124 NGDKKTSFLIGDGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDIN 183
Db 125 KNIITDTSLIGDGTGRGSANAQYFVFESEDERHFMHPHPEYSIITNIDFDHPDYFTGA 184
QY 184 DVDAFOEMAHNVKKGIIAWGDDEHLRKIEADVPYIYYGFKSDDIYAQNIQITDKGTAF 243
Db 185 DVNRAFNDYAKQVKKALFYVGGDDDELKKTIEAPAPIYVYGFEGNDPIAYDITRTTNGSDF 244
QY 244 DYVDGEFYDHLFSQYGDHTVNLAVLAIATSLVLEKLDVTNIKEALETFGGVKKRRNETT 303
Db 245 KYHQGEVIGQFHVPAVYKGNILNATAVIANLFAVAGIDMALVADHLKTFSGVKKRRFTEKI 304
QY 304 IANQVIVDDYAHHPREISATITETARKKYPHKEVAVFQPHTFSTRTOAFNLNEFAESLSKAD 363
Db 305 INDITIIDFAHPTETIVATIDARQKYSKEIVAFQPHTFSTRTOAFNLNEFAESLSKAD 364
QY 364 RVFLCEIFGSIRE-NTGALTIQDLIDK-IEGASLINESINVLQDFNAVILFMGAGDIQ 421
Db 365 SVYLAQIYGSAREVDKGEVREDLAAKIKPSQVTVENVSPLLDHDNAVYVFMGAGDIQ 424
QY 422 KIQNAVLDKL 431
Db 425 LYEHSEFEELL 434

RESULT 15

MURC_STRP8
ID MURC_STRP8 STANDARD; PRT; 442 AA.
AC Q8P2E1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-acetylmuramoyl-L-alanine synthetase).
GN MURC OR SPV18 0398.
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Sturdevant D.E., Ricklefs S.M., Porcella S.F., Sylva G.L., Barbard L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q., Parkins L.D., Dally J.A., Veasy L.G., Musser J.M.;
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18 group A Streptococcus strains associated with acute rheumatic fever outbreaks";
RT Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
CC -1- FUNCTION: Cell wall formation.
CC -1- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP + phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC -1- PATHWAY: Peptidoglycan biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: Belongs to the murCDEF family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AE005982; AAL97144.1; -.
HAMAP; MF_00046; -. 1.
InterPro; IPR000713; Mur_ligase.
InterPro; IPR004101; Mur_ligase_C.
InterPro; IPR005758; MurC.
Pfam; PF01225; Mur_ligase; 1.
Pfam; PF02875; Mur_ligase; 1.
TIGRFAMs; TIGR01082; murC; 1.
Ligase; ATP-binding; Cell wall;
Peptidoglycan synthesis; Complete proteome.
NP BIND 109 115 ATP (POTENTIAL).
SEQUENCE 442 AA; 49583 MW; 52535CCDDA3356E7 CRC64;

Very Match	49.9%	Score 1135;	DB 1;	Length 442;
1st Local Similarity	50.5%;	Pred. No. 1.5e-67;		
Matches 217;	Conservative	82;	Mismatches 129;	Indels 2;
Gap8	2;			

4 YHFGIKSGMSSLAQIMHDLGHEVQSGDIENYVFTEVALRNKGIRKILPFDANNIKEDMV 63
 |||:|||||||: ||: ||| |||||: |||: ||| |||||: |||: |||
 5 YHFGIKSGMSALALMIFOMGNRVQSGDVEKYFTQGLGEQAGITILPFSEDNITPME 64

```
64 VIQNAFASSEEIVRAHQKLDVVSYNDFLQIIDQYTSVAVTGAHGKTSTTGLLSHM 123
   : ||||| :::: | : : | : |||| : : | : ||||| : : ||||| : :
65 LTVGNAFRENKEVAYALRHQIPKRYHDFLGDFMKSFSISZAVAGHGKTSTTGLLSHV 124
```

124 NGDKKTSFLIGDTGMGIPESDYAFACEYRRHFI SYKPDYAIMTNIDFDHPDFFKDIN 183
|||:||||| : |||: |||: |||: |||: |||: |||: |||:
125 KNITDSYLIGDCTGRGSANAQVFYESDEYRFPYPHYEYSII TNIDFDHPDFFTGA 184

184 DVFDAPQEMAHNVKGIITAWGDDEHLRKIEADVPIYYGFKDSDDIYAQNIIQTDKGTAF 243
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
185 DVRNAFYDAKVKKALFYGGEDDELKKIELPAPIYYGFEGNDFIAYDITRTNGSDF 244

244 DYYVDGEFFYDHFUSPOYGHTVTNALAVIAISYLEKLDVNTIKEALETFGGVKRRFNETT 303
| | | | | : | | | | : | : | : | : | : | : |
245 KYKHOGEVIGOFHPAYGKNILNATVIANLFWAGIDMALVADHLKTFSGVKRRFTTEKI 304

304 IANQIVDDYAHHPREISATETARTKYPHKEVAVQPHTFSTRQAFLNEFAESLSKAD 363
:
305 VNGTIIIDFAHPTETIATDAAROKYPSIEVAIFOPHTFTRTIALEDFAALNEAD 364

```

364 RVFLCEFGSIRE-NTGALTIQDLIDK-IEGASLINEDSINVLFEQDFDNAVILFMFGAGDIQ 421
      -:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:|
365 SVYLAOTIGSSAREVDKGEVKVEDTAAKIIPSOVTVVENVSPLLDHNDNAVYVFMGAGDIQ 424

```

422 KLQAYLDKL 431
::: : |
425 LYHSFEELL 434

Search completed: June 3, 2004, 14:39:38
time : 19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 14:36:31 ; Search time 46 Seconds
(without alignments)
2997.423 Million cell updates/sec

Title: US-10-712-713-2

Perfect score: 2275

Sequence: 1 MTHYHFGVINGSGMSSIAQI.....GDIQLQNAVLDKLGKNAF 437

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

1: sp archaea:*

2: sp bacteria:*

3: sp fungi:*

4: sp human:*

5: sp invertebrate:*

6: sp mammal:*

7: sp mhc:*

8: sp organelle:*

9: sp phage:*

10: sp plant:*

11: sp rodent:*

12: sp virus:*

13: sp vertebrate:*

14: sp unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1438.5	63.2	445	16	Q833N6 enterococcu
2	1284.5	56.5	436	16	Q88WZ5 lactobacill
3	1163.5	51.1	443	16	Q8DY77 streptococc
4	1153.5	50.7	443	16	Q8E3U2 streptococc
5	1132.5	49.8	452	16	Q8DSF4 streptococc
6	633.5	27.8	486	16	Q999J1 clostridium
7	536.5	23.6	486	2	Q9RNM7
8	523.5	23.0	494	2	Q9F1M9
9	521.5	22.9	488	16	Q8E9P8 shewanella
10	506.5	22.3	468	16	Q7WFS3 bordetella
11	506.5	22.3	468	16	Q7W4B5 bordetella
12	506.5	22.3	468	16	Q7VUQ4 bordetella
13	502	22.1	486	16	Q87WY6 pseudomonas
14	500	22.0	465	16	Q83F17 coxiella bu
15	500	22.0	473	16	Q82VS2 nitrosomona
16	499.5	22.0	473	16	Q7V3P8 prochloroco

17	498.5	21.9	472	16	Q7VEJ1
18	498.5	21.9	475	16	Q7VMY1
19	490	21.5	487	2	Q9RGR6
20	490	21.5	491	16	Q83MF8
21	488.5	21.5	482	16	Q88N75 pseudomonas
22	484.5	21.3	512	16	Q8G4Q4 bifidobacte
23	480	21.1	464	16	Q82L07 streptomyce
24	474.5	20.9	467	16	Q89FU8 bradyrhizob
25	464	20.4	811	16	Q82LS4 chlamydophi
26	462.5	20.3	190	2	Q9KGV8
27	462	20.3	467	16	Q8A259 lactococcus
28	456	20.0	488	16	Q7V9C3 bacteroides
29	452.5	19.9	467	16	Q7UA71 synchococc
30	449	19.7	474	16	Q8KCE3 chlorobium
31	439.5	19.3	457	16	Q9CMA5
32	426	18.7	439	16	Q7VJ70
33	424	18.6	459	16	Q7VN73 haemophilus
34	420.5	18.5	461	16	Q8XVP8 ralstonia s
35	412.5	18.1	480	16	Q7VQI7 candidatus
36	405	17.8	484	16	Q8DE91 vibrio vuln
37	399	17.5	452	16	Q8PH24 xanthomonas
38	398	17.5	452	16	Q8P5P6 xanthomonas
39	393.5	17.3	475	16	Q83GN1 tropheryma
40	393.5	17.3	490	16	Q8F4J0 leptospira
41	392.5	17.3	475	16	Q83HK2 tropheryma
42	392	17.2	452	16	Q87SV8 vibrio para
43	390	17.1	458	16	Q9JUE5 neisseria m
44	383	16.8	458	16	Q9JRY9 neisseria m
45	382	16.8	465	16	Q8EYR2 leptospira

ALIGNMENTS

RESULT 1

ID	Q833N6	PRELIMINARY;	PRT;	445 AA.
AC	Q833N6;			
DT	01-JUN-2003 (TREMBLrel. 24, Created)			
DT	01-JUN-2003 (TREMBLrel. 24, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	UDP-N-acetylmuramate--alanine ligase.			
GN	MURC OR EFI908.			
OS	Enterococcus faecalis (Streptococcus faecalis).			
OC	Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.			
OX	NCBI_TaxID=1351;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=V583 / ATCC 700802;			
RX	MEDLINE=22550857; PubMed=12663927;			
RA	Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R., Gill S.R., Heidelberg J.F., Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Brinkac L., Beanan M., Tettelin H., Dodson R.J., Umayam L., Kolonay J., Madupu R., Nelson W., Daugherty S., DeBoy R.T., Durkin J., Hansen T., Shetty J., Khouri H., Vamathevan J., Tran B., Upton J., Ketchum K.A., Dougherty B.A., Fraser C.M., Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;			
RT	"Role of mobile DNA in the evolution of vancomycin-resistant Enterococcus faecalis.";			
RT	Science 299:2071-2074 (2003).			
RL	EMBL; AE016953; AAO81660.1; -.			
DR	TIGR; EFL908; -.			
DR	GO; GO:0005737; C:cytoplasm; IEA.			
DR	GO; GO:0005524; F:ATP binding; IEA.			
DR	GO; GO:0016874; F:ligase activity; IEA.			
DR	GO; GO:0008763; F:UDP-N-acetylmuramate-L-alanine ligase activity; IEA.			
DR	GO; GO:0009058; P:biosynthesis; IEA.			
DR	GO; GO:0009273; P:cell wall biosynthesis (sensu Bacteria); IEA.			
DR	InterPro; IPR005758; MurC.			
DR	InterPro; IPR000713; Mur_ligase.			
DR	InterPro; IPR004101; Mur_ligase_C.			
DR	Pfam; PF01225; Mur_ligase; 1.			
DR	Pfam; PF02875; Mur_ligase C; 1.			
DR	TIGRFAMs; TIGR01082; murC; 1.			

KW Ligase; Complete proteome.
SQ SEQUENCE 445 AA; 49999 MW; 12189004F4FB1E9A CRC64;
Query Match 63.2%; Score 1438.5; DB 16; Length 445;
Best Local Similarity 60.9%; Pred. No. 1.2e-82;
Matches 259; Conservative 79; Mismatches 86; Indels 1; Gaps 1;
QY 4 YHFGIKGSGMSSLAQIMHDLGHEVQSGSDIENYVTFEVALRNKGKILPFDANNIKEDMV 63
DB 9 YHFGIKGSGMSSLAQIMHDLGHEVQSGSDIENYVTFEVALRNKGKILPFDANNIKEDMV 68
QY 64 VIQGNAPASSHEEIVRAHQKLDVVSNDLFGQIIDQVTSVAVTGAHGTSTTGLLSHM 123
DB 69 VIAGNAPDTHETARATLGAELVIRHDFIARFTEPYTSIAVTGSHGKTSTTGLLAHVL 128
QY 124 NGDKKTSFLIGDGGMGMLPESDYFAFEACEYRRHFLSKPDYAIMTNIDDFDHPDYFKDIN 183
DB 129 SGINFTSYLIGDGGMGMLPESDYFAFEACEYRRHFLSKPDYAIMTNIDDFDHPDYFKDIN 188
QY 184 DVFDAPFOEMAHNVKKGIIAWGDDEHLRKIEADVPYIYVYGFKDSDDIYAQNIQITDKGTAF 243
DB 189 DVFSAPFMAHQVKGKIPAYGDKYLRQLESEVPYIYVYGVSEDDIQAQNIQITTEGSSF 248
QY 244 DVYVYDGEFDFHFLSPQYGDHVTNALAVIAISYLEKLDVTNIKEALETFGGVKKRRFNETT 303
DB 249 DVYHKDDFVGHEVLPFAFGHNNMAGLVIAVAYFEKLDQKVAEEMLSFGVKKRRFSEKK 308
QY 304 IANQIVDDVYAHHPREISATITETARKKYPHKEVAVFQPHTSRTQAFLNEFAESLSKAD 363
DB 309 VSDMIIVDDVYAHHPREISATITETARKKYPHKEVAVFQPHTSRTQAFLNEFAESLSKAD 368
QY 364 RVETCEIFGSRNTGALTIODLIDKIE-GASLINESDINVLEQFDNAVILFWAGAGDIQK 422
DB 369 EVFLCNIFGSRNTGALTIODLIDKIE-GASLINESDINVLEQFDNAVILFWAGAGDIQK 428
QY 423 LQNAV 427
DB 429 FEQAY 433

RESULT 2
Q88WZ5 PRELIMINARY; PRT; 436 AA.
ID Q88WZ5
AC Q88WZ5
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE UDP-N-acetylmuramate--alanine ligase (EC 6.3.2.8).
GN MURC OR LP 1462.
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIMB 8826 / WCFS1;
RX MEDLINE=24280296; PubMed=12566566;
RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
RA De Vos W.M., Siezen R.J.;
RT "Complete genome sequence of Lactobacillus plantarum WCFS1.";
RT Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
DR EMBL; AL935256; CAD63921.1; -;
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0008763; F:UDP-N-acetylmuramate-L-alanine ligase activity; IEA.
DR GO; GO:0003058; P:biogenesis; IEA.
DR InterPro; IPR000713; Mur_ligase.
DR InterPro; IPR004101; Mur_ligase.
DR Pfam; PF01225; Mur_ligase; 1.
DR Pfam; PF02875; Mur_ligase_C; 1.

KW Ligase; Complete proteome.
SQ SEQUENCE 436 AA; 48750 MW; 25A582D850238CD4 CRC64;
Query Match 56.5%; Score 1284.5; DB 16; Length 436;
Best Local Similarity 55.7%; Pred. No. 7.6e-82;
Matches 240; Conservative 75; Mismatches 115; Indels 1; Gaps 1;
QY 2 THYFVGKSGMSSLAQIMHDLGHEVQSGSDIENYVTFEVALRNKGKILPFDANNIKED 61
DB 5 TVYFVGKSGMSSLAQIMHDLGHEVQSGSDIENYVTFEVALRNKGKILPFDANNIKED 64
QY 62 MVYIQGNAPASSHEEIVRAHQKLDVVSNDLFGQIIDQVTSVAVTGAHGTSTTGLLSH 121
DB 65 LTVIAGNSFTDHPETKKAEMGLPYRYRHEFLGKLMGEGTSGVAGTHGKTSTTGLLSH 124
QY 122 VMNGDKKTSFLIGDGGMGMLPESDYFAFEACEYRRHFLSKPDYAIMTNIDDFDHPDYFKD 181
DB 125 VUSHTAPTSLYIGDGGMGMLPESDYFAFEACEYRRHFLSKPDYAIMTNIDDFDHPDYFKD 184
QY 182 INDVFDAPFOEMAHNVKKGIIAWGDDEHLRKIEADVPYIYVYGFKDSDDIYAQNIQITDKGT 241
DB 185 LADVQSAFQFGNQVKGKIPAYGDKYLRQLESEVPYIYVYGVSEDDIQAQNIQITTEGSSF 244
QY 242 DVYVYDGEFDFHFLSPQYGDHVTNALAVIAISYLEKLDVTNIKEALETFGGVKKRRFNETT 301
DB 245 SPEVKYHDESGLGKFEIPLFGEHNNMAGLVIAVAYFEKLDQKVAEEMLSFGVKKRRFSE 304
QY 302 TTIANQIVDDVYAHHPREISATITETARKKYPHKEVAVFQPHTSRTQAFLNEFAESLSK 361
DB 305 HQVGMVMIIVDDVYAHHPREISATITETARKKYPHKEVAVFQPHTSRTQAFLNEFAESLSK 364
QY 362 ADVFLCEIFGSRNTGALTIODLIDKIE-GASLINESDINVLEQFDNAVILFWAGAGDI 420
DB 365 ADHVLTNIFGSRNTGALTIODLIDKIE-GASLINESDINVLEQFDNAVILFWAGAGDI 424
QY 421 QKQNAVYDKL 431
DB 425 QKQNAVYDKL 435

RESULT 3
Q8DY77 PRELIMINARY; PRT; 443 AA.
ID Q8DY77
AC Q8DY77
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE UDP-N-acetylmuramate--alanine ligase.
GN MURC OR SAGI615.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547;
RA Tettelin H., Maignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Bean M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
RA Rinaldo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
DR EMBL; AB014265; AAN00479.1; -;
DR TIGR; SAGI615; -;
DR GO; GO:0005840; C:ribosome; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.


```
DR GO: GO:0016874; F:ligase activity; IEA.
DR GO: GO:0003735; F:structural constituent of ribosome; IEA.
DR GO: GO:0008763; F:UDP-N-acetylmuramate-L-alanine ligase activity; IEA.
DR GO: GO:0009058; P:biosynthesis; IEA.
DR GO: GO:0009273; P:cell wall biosynthesis (sensu Bacteria); IEA.
DR GO: GO:0006412; P:protein biosynthesis; IEA.
DR InterPro: IPR005758; MurC.
DR InterPro: IPR000713; Mur_ligase.
DR InterPro: IPR004101; Mur_ligase_C.
DR InterPro: IPR001706; Ribosomal_L35.
DR Pfam: PF01225; Mur_ligase; 1.
DR TIGRFAMs: TIGR01082; murC; 1.
DR PROSITE: PS00936; RIBOSOMAL_L35; 1.
KW Ligase; Complete proteome.
SQ SEQUENCE 443 AA; 49852 MW; C1D6FC7DE6017228 CRC64;

Query Match 51.1%; Score 1163.5; DB 16; Length 443;
Best Local Similarity 51.3%; Pred. No. 2.4e-73;
Matches 223; Conservative 81; Mismatches 120; Indels 11; Gaps 4;

QY 4 YHFGIKSGMSLSLAQIMHDLGHEVQSGDIENYVTFEVALRNKGKILPFDANNIKEDMV 63
DB |||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
5 YHFIGIKSGMSALALMLHQMGNVQSGDVKYFTQGLEQAGVTILPFPSPNNISEDLE 64
64 VIQGNAP-ASSHEIIVRAHQKLDVVSYNDFLQIIDDQVTSVAVTGAGHKSTTGL 118
65 ILAGNAFRPDNNEELAYVIEKGQFK----RVHEFLGDFMRQTSIGVAGAGHKSTTGL 120
119 LSHVNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMNTIDFDPDY 178
121 LAHVLKNITDTSFLIGDGTGRGSANAYVFVEADERHFMFVHPHYSIITNIDFDPDY 180
179 FKDINDVFADQEMAHNVKGIAMGDDHRLKRIEADVPYIYVYFGKSDDIYAQNTQITD 238
181 FTGLEDFVNAFNDAKQVQKGLFIYGEDPKLHEITSEAPIYIYVYFGSDNDFIAKDIRTV 240
239 KGTAFDVYDGFYDHFSLSPQGDHVTNALAVIALSYLEKLDVTNIKEALETFPGVKRR 298
241 NGSDFKVFNQSEIQGFHVPAYGKHNILNATAVIANLYIMGIDMALVAEHLKTFPGVKRR 300
299 FNETTIANQVIVDDVAHHPREISATITETARKKYPHKEVVAVFQPHFSTRTOAFLENAFAS 358
301 FTEKLIIDTVIIDDFAHHPTEIATDAAKQYPSKEIVAFQPHFTFTRIALLDFAHLSOA 360
359 LSKADRVFLCEIFGSIRES-NTGALTIQDLIDKI-EGASLINESINVLQFDNAVILFMG 416
361 LSQADSVYLAQIYGSAREVNGEVEDLAAKIVKHSDLVTVENVSPLLNHDNAVYVFMG 420
417 AGDIQKQNAVLDKL 431
421 AGDIQLYERSFEELL 435

RESULT 4
ID Q8E3U2 PRELIMINARY; PRT; 443 AA.
AC Q8E3U2;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN MURC OR GBS1664.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MEM316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Kusniok C., Buchrieser C., Chevalier F., Prangeul L.,
RA Meadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
```

```
RA "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease.";
RL Mol. Microbiol. 45:1499-1513 (2002).
DR EMBL; AL766852; CAD47323.1; -.
DR Sgaliast; gbs1664; -.
DR GO: GO:0005840; C:ribosome; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016874; F:ligase activity; IEA.
DR GO: GO:0003735; F:structural constituent of ribosome; IEA.
DR GO: GO:0008763; F:UDP-N-acetylmuramate-L-alanine ligase activity; IEA.
DR GO: GO:0009058; P:biosynthesis; IEA.
DR GO: GO:0009273; P:cell wall biosynthesis (sensu Bacteria); IEA.
DR GO: GO:0006412; P:protein biosynthesis; IEA.
DR InterPro: IPR005758; MurC.
DR InterPro: IPR000713; Mur_ligase.
DR InterPro: IPR004101; Mur_ligase_C.
DR InterPro: IPR001706; Ribosomal_L35.
DR Pfam: PF01225; Mur_ligase; 1.
DR TIGRFAMs: TIGR01082; murC; 1.
DR PROSITE: PS00936; RIBOSOMAL_L35; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 443 AA; 49836 MW; 0BDBD0CDB62BBD9 CRC64;

Query Match 50.7%; Score 1153.5; DB 16; Length 443;
Best Local Similarity 51.0%; Pred. No. 1.2e-72;
Matches 220; Conservative 82; Mismatches 126; Indels 3; Gaps 3;

QY 4 YHFGIKSGMSLSLAQIMHDLGHEVQSGDIENYVTFEVALRNKGKILPFDANNIKEDMV 63
DB |||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:
5 YHFIGIKSGMSALALMLHQMGNVQSGDVKYFTQGLEQAGVTILPFPSPNNISEDLE 64
64 VIQGNAP-ASSHEIIVRAHQKLDVVSYNDFLQIIDDQVTSVAVTGAGHKSTTGLSHV 122
65 ILAGNAFRPDNNEELAYVIEKGHFKRYHFEFLGDFMRQTSIGVAGAGHKSTTGLLAHV 124
123 MNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMNTIDFDPDYFKDI 182
125 LKNITDTSFLIGDGTGRGSANAYVFVEADERHFMFVHPHYSIITNIDFDPDYFTGL 184
183 NDVFADQEMAHNVKGIAMGDDHRLKRIEADVPYIYVYFGKSDDIYAQNTQITDKGTA 242
185 EDVFNAFNDYAKQVQKGLFIYGEDSKLHEITSKAPIYIYVYFGSDNDFIAKDIRTVNGSD 244
243 FDVYDGFYDHFSLSPQGDHVTNALAVIALSYLEKLDVTNIKEALETFPGVKRRFNET 302
245 FKVFYNOEIEIQGFHVPAYGKHNILNATAVIANLYIMGIDMALVAEHLKTFSGVKRRFTEK 304
303 TIANQVIVDDVAHHPREISATITETARKKYPHKEVVAVFQPHFSTRTOAFLENAFASLSKA 362
305 IIDDTVIIDDFAHHPTEIATDAAKQYPSKEIVAFQPHFTFTRIALLDFAHLSOA 364
363 DRVFLCEIFGSIRES-NTGALTIQDLIDKI-EGASLINESINVLQFDNAVILFMGAGDI 420
365 DSVYLAQIYGSAREVNGEVEDLAAKIVKHSDLVTVENVSPLLNHDNAVYVFMGAGDI 424
421 QKQNAVLDKL 431
425 QLYERSFEELL 435

RESULT 5
Q8DSP4
ID Q8DSP4 PRELIMINARY; PRT; 452 AA.
AC Q8DSP4;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative UDP-N-acetyl muramate-alanine ligase (EC 6.3.2.8).
GN MURC OR SMU.1731.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
```


DE UDP-N-acetylmuramate--alanine ligase.
GN MURC OR S04218.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=MR-1;
RC MEDLINE=22297686; PubMed=12368813;
RX Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Trapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Neallson K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis";
RL Nat. Biotechnol. 20:1118-1123 (2002).
RL EMBL; AE015855; AAN57190.1; -;
DR TIGR; S04218; -;
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0008763; F:UDP-N-acetylmuramate-L-alanine ligase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR GO; GO:0009273; P:cell wall biosynthesis (sensu Bacteria); IEA.
DR InterPro; IPR005758; MurC.
DR InterPro; IPR000713; Mur_ligase.
DR InterPro; IPR004101; Mur_ligase_C.
DR Pfam; PF01225; Mur_ligase; 1.
DR Pfam; PF02875; Mur_ligase_C; 1.
DR TIGRFAMS; TIGR01082; murC; 1.
DR Ligase; Complete proteome.
KW SEQUENCE 488 AA; 53168 MW; DC45914D4869D353 CRC64;
SQ
Query Match 22.9%; Score 521.5; DB 16; Length 488;
Best Local Similarity 32.6%; Pred. No. 3e-28;
Matches 146; Conservative 87; Mismatches 182; Indels 33; Gaps 13;
Qy 5 HFVIGKSGMSLAQIMHDLGHEVQSGDI-ENYVTFEVALRNKGIKI-LPFDANNIKEDM 62
Db 24 HFVIGGAGMGIAEVLNVEGVSSGSDIAQNAVTDRLCL--LGAKIHIGHGADNVQOQD 81
Qy 63 VVIQGNAPASSHEEIVRAHQLKLDVSNDFLQGIIDQYTSVAVTGAGHKTSITGLLSHV 122
Db 82 VVVVSTAINPONPELIAAKEIRIPVRAEMLAELMYRHGVAITAGTKTTISLASL 141
Qy 123 M-NGDKKTSFLIGD-----GTGMGLPESDYFAFEACEYRHRFLSKYKDYAIAWNTIDFHP 176
Db 142 YQAGRDPTFVIGLLNSAGTNARLGTSKYLAEADESDASFLHLQPMVSVVNTIEADHM 201
Qy 177 D-YFKDINDVDFAPQEAHNVK-KGI-TAWGDDEHLRIEADVP--IYYGFKSDSDIYA 231
Db 202 DTYGDFEKLKSTFVDVFLHNLFFYGVAVVVCIDDPVVRIMPRISHIVTYGFRDDAVQA 261
Qy 232 QNIQITDKGTAFDVYDGEFFYDHELS PQYGDHVTNLALAVIAISYLEKLDVNIKEALET 291
Db 262 LNFSSQGHQCRFTVRKKGEDLDLNLNLPQGNVNLALAAIAVATEDEIDDSAILQALAE 321
Qy 292 FGGVKKRRFNE-----TTIANQVIVDDYAHHPREISATITETARKKYPHKVEAVPQFTFS 346
Db 322 FQIGRRFOHLKFKATPKGEVMLVDDYGHHPSEVAATIKAAAGWPEKELVWAYQPHRYT 381
Qy 347 RTQAFINEFAESLSKADRVFLCEFGS-----IRENTGALTIQDLIDKIEGAS--L 395
Db 382 RTDRDYEDFIEVLSQVDCLLLDVYSAGRAPIPGADGRALCRISIRLQGLDPIFIASPEQ 441
Qy 396 INEDSINVLEQFDNAVILFMGAGDIQKL 423
Db 442 LAEVLFPDLVQEGD--LLLTQAGNIGAL 467

```
RESULT 11
Q7W4B5 PRELIMINARY; PRT; 468 AA.
AC Q7W4B5
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE UDP-N-acetylmuramate--alanine ligase (EC 6.3.2.8).
GN MURC OR BP3751
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaibia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
RA "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640434; CAE39034.1; -.
KW Ligase; Complete proteome.
SQ SEQUENCE 468 AA; 49404 MW; 045E233D594FECFD CRC64;

Query Match 22.3%; Score 506.5; DB 16; Length 468;
Best Local Similarity 31.7%; Pred. No. 3.2e-27;
Matches 145; Conservative 75; Mismatches 195; Indels 43; Gaps 14;

QY 3 HYHFGVIGKSGMSSLAQIMHDHGHVEQSGDIENYVTEVALRNKGKI-LPFDANNIKED 61
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 7 HIHFVGVGSGMSSLAQIMHDHGHVEQSGDIENYVTEVALRNKGKI-LPFDANNIKED 65
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 62 MVVIOGNFASHEEIVRAHQLKLDVSYNDPLGQIIDQYTSVAVTCAGKTKSTTGLLGH 121
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 66 GAIVTSTAVAGNPEVLAARAARIPVPRVAVMLAELMRKRGIAVAGTGKTTTSLVAS 125
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 122 VM-NGDKKTSFLIGD-----GTGMLPESDYFAFEACEYRRHFLSKYKPDYAIMTIDFH 175
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 126 VLAAGGLDPTFVIGGLTSGANARLGQGEYIVVEADESDASFLNLLPVAIVTNIDAH 185
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 176 PD-YFKDINDVDFADQEMAHNVK--KGIANGDDEHLRKEADV--PIYYGFKDSDDTY 230
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 186 MDTYGHVVARLKSATFETQRLPFYGSAILCADDANVREIMPFVSRPITTYGLSPDAQVC 245
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 231 AQNIQITDKGTFADVVVDGEFFYHFLSPQY-----GDHTVLNALAVIAISYLEKLDVTN 284
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 246 AODVQA--DGTMRFTVQRDRDVLVPALQVELNLPGLHNVNRNALAAIAVATELGVDAA 303
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 285 IKEALETFGGVKRRFNE-----TTIANQ-----VIVDDYAHHPREISATITARKKYPHKE 335
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 304 IREALAFAKVGRRFTQWGDLPVPAAHGGGFTFLVDYDGHHPVEMAATLAAARGANPQRR 363
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 336 VVAVFQPHFTSRTOAFNLNFAESLSKADRVFLCEIFGS-----IRENTGALTIQDL 386
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 364 IVLAFQPHRYTRTRDCFDVFRVLGSADGVLLTEVYAAEAPLVAADGRALSRALRVAGK 423
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 387 IDK--IEGASLINEDSINVLEQFDNAVILFMAGDIQK 422
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 424 VEPFVEDVGELPQAILDFVR--DGDVVVMVAGGSISK 459
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 12
Q7W4B5 PRELIMINARY; PRT; 486 AA.
AC Q7W4B5
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE UDP-N-acetylmuramate--alanine ligase (EC 6.3.2.8).
GN MURC OR BP3751
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaibia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
RA "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640420; CAE43293.1; -.
KW Ligase; Complete proteome.
SQ SEQUENCE 468 AA; 49416 MW; 1E77233D436GE724 CRC64;
```



```
QY 429 DKLG 432
Db 462 QTHG 465

RESULT 15
Q82VS2 PRELIMINARY; PRT; 473 AA.
ID Q82VS2
AC Q82VS2, 2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE MurC; UDP-N-acetylmuramate--alanine ligase protein (EC 6.3.2.8).
GN MURC OR NE0992.
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OX NCBI_TaxID=915;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19718 /IFO 14298;
RX MEDLINE=22586410; PubMed=12700255;
RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.,
RA Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
RA Arciero D.M., Hommes N.G., Whittaker M.M., Arp D.J.;
RA "Complete genome sequence of the ammonia-oxidizing bacterium and
RT obligate chemolithoautotroph Nitrosomonas europaea.";
RL J. Bacteriol. 185:2759-2773 (2003).
DR EMBL; BX321859; CAD84903.1; -.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0008763; F:UDP-N-acetylmuramate-L-alanine ligase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR GO; GO:0009273; P:cell wall biosynthesis (sensu Bacteria); IEA.
DR InterPro; IPR005758; MurC.
DR InterPro; IPR000713; Mur_ligase.
DR InterPro; IPR004101; Mur_ligase_C.
DR Pfam; PF01225; Mur_ligase_1.
DR Pfam; PF02875; Mur_ligase_C; 1.
DR TIGRFAMs; TIGR01082; murC; 1.
KW Ligase; Complete proteome.
SQ SEQUENCE 473 AA; 51271 MW; BE5A7F7C42471FD7 CRC64;

Query Match 22.0%; Score 500; DB 16; Length 473;
Best Local Similarity 32.2%; Pred. No. 9.3e-27;
Matches 151; Conservative 74; Mismatches 176; Indels 68; Gaps 18;

QY 3 HYHFGVIGKSGMSLAQIMHDLGHEVQGSIENTVTFTEVALRNKGIKI-LPFDANNIKED 61
Db 7 HIHFVGIGSGMGIAEVLINLGFQISGSDMHSNSTTR-LQCLGAVIHTHAAENIQSA 65
QY 62 MVVIOGNAPASSHEEIVRAHQKLDVVSYNDFLGGIIDQYTSVAVTGAHGKTSITGLLSH 121
Db 66 DAVVISTAHSDNPEVIAAERRIPVVPRAWMLAEHLRLRRGIALAGTHGKTTTSLVAS 125
QY 122 VM-NGDKTSLFLIGD-----GTGMLPESDYFAFEACEYRRHFLSKYKPDYAIMTNIDFDH 175
Db 126 ILAAGQDPTFVIGGKLKTVDSHARLKGGEFVVVEADESDASFLYLQPLVLTVTNIDADH 185
QY 176 -PDYFKDINDVFAPQF-MAHNVKKGI-IAWGDDHLRKIEADV--PIYYGFKDSD-DI 229
Db 186 MSTYEHDFNRLKQTFVEFIEHLPFGYMAVLCDVDPHVREIISMITRPVTTYGIASEDAQI 245
QY 230 YACNIQITDKGTAFDVVDGEFYD---HEL-----SPOY-----GDHTVLNALA 270
Db 246 CATNIR-----HDCRMHFLAHGVNGSPRTLEVTNLPGKHVNLNALA 289
QY 271 VIAISVLEKLDVTN--IKEALETFGGVKRRF---NETTIANQ---VIVDDVYAHHPREISA 322
Db 290 AIAVG--NELGVDEALVKALATFGGVDRRFQQYGEIPLPDQGSFALIDDYGHHPAEIAA 347
QY 323 TITARKKYPHKEVAVFQPFTHSRTOAFLNEFAESLSKADRVFLCFIGSIRENTGALT 382
```

```
Db 348 TWAARNAPFGRRLVLAFOPHRYSTRDLFPEDFVRVLUGADVLLLTTEVYPAGEPPIIAAD 407
QY 383 IQDLIDKIEGASLINESINVLQF-----DNAVILFMGAGDIQK 422
Db 408 SKSLABAIRVQGKIETPIYIEQIDELKATIHTIAQDGDVILIMGAGSICK 456
```

Search completed: June 3, 2004, 14:40:38
Job time : 48 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2004, 22:02:04 ; Search time 5350 Seconds
(without alignments)
10677.779 Million cell updates/sec

Title: US-10-712-713-1

Perfect score: 1318

Sequence: 1 atgacacactatcattttgt.....gaaaaatgcgttttaagctt 1318

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_btg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_man.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1314	99.7	1335	6	AX742069	AX742069 Sequence
2	1314	99.7	2424	6	AR354274	AR354274 Sequence
3	1312.4	99.6	1335	6	AX742071	AX742071 Sequence
C	1311	99.5	1332	6	AX617936	AX617936 Sequence
C	1310.8	99.5	301550	1	AP003134	AP003134 Staphyloc
C	1310.8	99.5	342600	1	AP003363	AP003363 Staphyloc
C	1309.2	99.3	272850	1	AP004828	AP004828 Staphyloc
C	1302.8	98.8	1314	1	AF034076	AF034076 Staphyloc
C	1302.8	98.8	1314	1	AX191741	AX191741 Sequence
C	1293.2	98.1	1351	6	E35613	E35613 MurC. 6/200
C	1293.2	98.1	1351	6	BD178030	BD178030 MurC. 4/2
C	918.8	69.7	300029	1	AE016748	AE016748 Staphyloc
C	637.8	48.4	660	6	E35614	E35614 MurC. 6/200
C	637.8	48.4	660	6	BD178031	BD178031 MurC. 4/2
C	597.6	45.3	291804	1	AE017039	AE017039 Bacillus
C	596	45.2	300375	1	AE017013	AE017013 Bacillus
C	570.4	43.3	619	6	AR194630	AR194630 Sequence
C	553.4	42.0	302241	1	AE016953	AE016953 Enterococ
C	545.2	41.4	299950	1	AP004600	AP004600 Oceanobac
C	523	39.7	239050	1	AL596169	AL596169 Listeria
C	523	39.7	349980	6	AX417045	AX417045 Sequence
C	517.4	39.3	349980	6	AX413017	AX413017 Sequence
C	516.4	39.2	240050	1	AL591979	AL591979 Listeria
C	516.4	39.2	349980	6	AX641669	AX641669 Sequence
C	500.8	38.0	199322	1	BSUB0016	Z59119 Bacillus su
C	500.8	38.0	220060	1	AF008220	AF008220 Bacillus
C	495.6	37.6	303250	1	AP001518	AP001518 Bacillus
C	456	34.6	10417	1	AE006433	AE006433 Lactococc
C	450.8	34.2	1332	6	AX607160	AX607160 Sequence
C	450.8	34.2	20501	1	AE014265	AE014265 Streptoco
C	450.8	34.2	44145	6	AX602195	AX602195 Sequence
C	450.8	34.2	174050	1	SAG766852	AL766852 Streptoco
C	447.8	34.0	10470	1	AE006499	AE006499 Streptoco
C	446.2	33.9	52276	1	AE014141	AE014141 Streptoco
C	446.2	33.9	323825	1	AP005146	AP005146 Streptoco
C	443	33.6	11500	1	AE009982	AE009982 Streptoco
C	434.8	33.0	302050	1	AL935256	AL935256 Lactobaci
C	426.2	32.3	12221	1	AE015001	AE015001 Streptoco
C	412.4	31.3	2320	1	BACUNAM	I31845 Bacillus su
C	402.4	30.5	10461	1	AE008507	AE008507 Streptoco
C	400.8	30.4	1332	6	AX569902	AX569902 Sequence
C	400.8	30.4	10344	1	AE007447	AE007447 Streptoco
C	400.8	30.4	11864	6	AR218829	AR218829 Sequence
C	400.8	30.4	11864	6	BD003741	BD003741 Polynucle
C	400.8	30.4	297172	2	SPNEU1905	AL449927 Streptoco

ALIGNMENTS

RESULT 1	AX742069	Sequence 27 from Patent WO03025007.	1335 bp	DNA	linear	PAT 10-MAY-2003
LOCUS	AX742069	Sequence 27 from Patent WO03025007.				
DEFINITION	AX742069	Sequence 27 from Patent WO03025007.				
ACCESSION	AX742069	Sequence 27 from Patent WO03025007.				
VERSION	AX742069.1	GI:30524566				
KEYWORDS		Staphylococcus aureus				
SOURCE		Staphylococcus aureus				
ORGANISM		Bacteria; Firmicutes; Bacillales; Staphylococcus.				
REFERENCE		1				
AUTHORS		Edwards, A., Dharamsi, A., Vedadi, M., Alam, M. Z., Awrey, D., Beattie, B., Domagala, M., Houston, S., Kanagarajah, D., Nethery, K., Ng, I., Mansoury, K., McDonald, M. L., Pinder, B., Viola, C. and Wrezel, O.				

TITLE Novel purified polypeptides involved in membrane biosynthesis
JOURNAL Patent: WO 03025007-A 27 27-MAR-2003;
Affinium Pharmaceuticals, Inc. (CA)
FEATURES Location/Qualifiers
1..1335
/organism="Staphylococcus aureus"
/mol_type="genomic DNA"
/db_xref="taxon:1280"

ORIGIN

Query Match	99.7%	Score 1314	DB 6	Length 1335
Best Local Similarity	100.0%	Pred. No. 5.5e-190		
Matches 1314	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	ATGCACACATCATCTTTGTCGGAAATTAAGGTTCTGGCATGAGTTCATTTAGCACAATC	60	
DB	22	ATGCACACATCATCTTTGTCGGAAATTAAGGTTCTGGCATGAGTTCATTTAGCACAATC	81	
QY	61	ATGCATGATTTAGGCACATGAAGTTCAAGGATCGGATATTCAGAACTACGTTATTTACAGAA	120	
DB	82	ATGCATGATTTAGGCACATGAAGTTCAAGGATCGGATATTCAGAACTACGTTATTTACAGAA	141	
QY	121	GTTCGCTTTAGAAATAAGGGGATAAAAATATATACCATTTGATGCTAATAATACATAAAGAA	180	
DB	142	GTTCGCTTTAGAAATAAGGGGATAAAAATATATACCATTTGATGCTAATAATACATAAAGAA	201	
QY	181	GATATGCTAGCTATACAAAGCTAATGCATTCGCGAGTAGCCATGAGAAATAGTACGTGCA	240	
DB	202	GATATGCTAGCTATACAAAGCTAATGCATTCGCGAGTAGCCATGAGAAATAGTACGTGCA	261	
QY	241	CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTTAGGCACAGATTTATGATCAA	300	
DB	262	CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTTAGGCACAGATTTATGATCAA	321	
QY	301	TATACCTTCAGTACTGTAACTGGTGACATGGTAAACCTTCTCAACAGGTTATTTATCA	360	
DB	322	TATACCTTCAGTACTGTAACTGGTGACATGGTAAACCTTCTCAACAGGTTATTTATCA	381	
QY	361	CATGTTATGAATGCTGATAAAAGACTTCATTTTTAATTCGTCGATGCGACAGGATGCGGA	420	
DB	382	CATGTTATGAATGCTGATAAAAGACTTCATTTTTAATTCGTCGATGCGACAGGATGCGGA	441	
QY	421	TTGCCTGAAAGTGATTTATTTTCGCTTTTGAGGCATGTGAAATATAGACGTCACTTTTTAAAGT	480	
DB	442	TTGCCTGAAAGTGATTTATTTTCGCTTTTGAGGCATGTGAAATATAGACGTCACTTTTTAAAGT	501	
QY	481	TATAAACCTTGATTTACGCAATATATGCAAAATATTGATTTTCGATCATCTGATTTATTTAAA	540	
DB	502	TATAAACCTTGATTTACGCAATATATGCAAAATATTGATTTTCGATCATCTGATTTATTTAAA	561	
QY	541	GATATTAAATGATGTTTTTGATGCATTTCCAGAAATGCGACATAATGTTAAAAAGGTATT	600	
DB	562	GATATTAAATGATGTTTTTGATGCATTTCCAGAAATGCGACATAATGTTAAAAAGGTATT	621	
QY	601	ATTGCTTTGGGGTGATGATGAACATCTACGTPAAAAATTTGAAGCAGATGTTCCAAATTTATTAT	660	
DB	622	ATTGCTTTGGGGTGATGATGAACATCTACGTPAAAAATTTGAAGCAGATGTTCCAAATTTATTAT	681	
QY	661	TATGGATTTAAAGATTCGGANGACATTTATGCTCAAAATATTCAAATTCGGAATAAGGT	720	
DB	682	TATGGATTTAAAGATTCGGANGACATTTATGCTCAAAATATTCAAATTCGGAATAAGGT	741	
QY	721	ACTGCTTTTGATGCTGATGTTGGATGGTGAGTCTTTATGATGATCACTTCCTGCTCCACATAT	780	
DB	742	ACTGCTTTTGATGCTGATGTTGGATGGTGAGTCTTTATGATGATCACTTCCTGCTCCACATAT	801	
QY	781	GGTGACCATACAGTTTTAAATGCAATTAGCTGTAAATTCGGAATTAGTTATTTAGAGAAGCTA	840	
DB	802	GGTGACCATACAGTTTTAAATGCAATTAGCTGTAAATTCGGAATTAGTTATTTAGAGAAGCTA	861	
QY	841	GATGTTTCAAATATTAAGAAGCATTAGAAACGTTTTGGTGGTGTAAACGTCGTTCAAT	900	
DB	862	GATGTTTCAAATATTAAGAAGCATTAGAAACGTTTTGGTGGTGTAAACGTCGTTCAAT	921	

901	QY	GAACCTACAAATTCGCAAACTCAAGTTATTGTAGATGATTATGCACACCAATCCAGAGAAATT	960
922	Db	GAAACTACAAATTCGCAAACTCAAGTTATTGTAGATGATTATGCACACCAATCCAGAGAAATT	981
961	QY	AGTGCTACAAATTCGCAAACTCAAGTTATTGTAGATGATTATGCACACCAATCCAGAGAAATT	1020
982	Db	AGTGCTACAAATTCGCAAACTCAAGTTATTGTAGATGATTATGCACACCAATCCAGAGAAATT	1041
1021	QY	CAACCCACACACTTTCTCTAGAACACAGGCCATTTTTAAATGAAATTCGAGAAAAGTTTAAGT	1080
1042	Db	CAACCCACACACTTTCTCTAGAACACAGGCCATTTTTAAATGAAATTCGAGAAAAGTTTAAGT	1101
1081	QY	AAAGCAGATCGTGATTCCTTATGTGAAATTTTTGGATCAATTTAGAGAAAATATCTGGCGCA	1140
1102	Db	AAAGCAGATCGTGATTCCTTATGTGAAATTTTTGGATCAATTTAGAGAAAATATCTGGCGCA	1161
1141	QY	TTAAACGATCAAGATTTAAATGTGATAAAATTTGAGGTGCATCGTTAAATTAATGAGATTCT	1200
1162	Db	TTAAACGATCAAGATTTAAATGTGATAAAATTTGAGGTGCATCGTTAAATTAATGAGATTCT	1221
1201	QY	ATTTAATGTTAGACAAATTTGATAATTCGTTATTTTATTATGCGTGACAGGTGATTT	1260
1222	Db	ATTTAATGTTAGACAAATTTGATAATTCGTTATTTTATTATGCGTGACAGGTGATTT	1281
1261	QY	CAAAAAATTCAAAATGCATATTTAGATAAATTTAGGCATGAAAAATCGTTTTAA	1314
1282	Db	CAAAAAATTCAAAATGCATATTTAGATAAATTTAGGCATGAAAAATCGTTTTAA	1335

RESULT 2
 AR354274
 LOCUS
 DEFINITION
 Sequence 392 from patent US 6593114.
 AR354274
 ACCESSION
 AR354274.1 GI:33760358
 KEYWORDS
 .
 SOURCE
 Unknown.
 ORGANISM
 Unknown.
 REFERENCE
 1 (bases 1 to 2424)
 AUTHORS
 Kunsch,C.A., Choi,G.H., Barash,S., Dillon,P.J., Fannon,M.R. and
 Rosen,C.A.
 TITLE
 Staphylococcus aureus polynucleotides and sequences
 JOURNAL
 Patent: US 6593114-A 392 15-JUL-2003;
 FEATURES
 Location/Qualifiers
 1..2424
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN

ORIGIN		/ mol_type= genomic DNA	
Query Match	99.7%	Score 1314;	DB 6; Length 2424;
Best Local Similarity	100.0%;	Pred. No. 4.6e-190;	
Matches 1314; Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	ATGACACACATCATTTTGTTCGGAAATTAAGGTTCTGGCATGAGTTCATTAGCACAAATC	60
DB	627	ATGACACACATCATTTTGTTCGGAAATTAAGGTTCTGGCATGAGTTCATTAGCACAAATC	686
QY	61	ATGCATGATTTTAGCACATGAAGTTTCAAGGATCGGATATTGAGAACTACGTTATTTACAGAA	120
DB	687	ATGCATGATTTTAGCACATGAAGTTTCAAGGATCGGATATTGAGAACTACGTTATTTACAGAA	745
QY	121	GTTCGCTCTTAGAAATAAGGGGATAAAAAATTAACATTTGATGCTCTAATAACATAAAGAA	180
DB	747	GTTCGCTCTTAGAAATAAGGGGATAAAAAATTAACATTTGATGCTCTAATAACATAAAGAA	806
QY	181	GATATCGGTAGTTATACAAGSTAATGCATTCGCGAGTAGCCATCAAGAAATAGTACGTGCA	240
DB	807	GATATCGGTAGTTATACAAGSTAATGCATTCGCGAGTAGCCATCAAGAAATAGTACGTGCA	866
QY	241	CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTTTAGGACAGATTTATTGATCAA	300

867	DB	CATCAATTCGAAATTAGATGTTGTAAAGTTATATAGATTTTTTTAGGACAGATTAATTGATCAAA	926
301	QY	TATACTTCAGTAGCTGTAACCTGGTGCACATGGGTAAACACTTCTACAAACAGGTTTATTATATCA	360
927	DB	TATACTTCAGTAGCTGTAACCTGGTGCACATGGGTAAACACTTCTACAAACAGGTTTATTATCA	986
361	QY	CATGTTATGAATGGTGATAAAAAGACTTCATTTTTTAAATTTGGTGATGGCACAAGTATGGGA	420
987	DB	CATGTTATGAATGGTGATAAAAAGACTTCATTTTTTAAATTTGGTGATGGCACAAGTATGGGA	1046
421	QY	TTGGCTCGAAAGTGATTAATTTTCGCTTTTGAGGCATGTGAATATAGACGTCACATTTTTAACT	480
1047	DB	TTGGCTCGAAAGTGATTAATTTTCGCTTTTGAGGCATGTGAATATAGACGTCACATTTTTAACT	1106
481	QY	TATAAACCTGATTACGCCAATATTATGACAAAATATTGATTTTGGATCATCTCATTAATTTTAAA	540
1107	DB	TATAAACCTGATTACGCCAATATTATGACAAAATATTGATTTTGGATCATCTCATTAATTTTAAA	1166
541	QY	GATATTAATGATGTTTTGATGCAATTTCCAAGAAATGGCACAATAATGTTTAAAAAGGTATT	600
1167	DB	GATATTAATGATGTTTTGATGCAATTTCCAAGAAATGGCACAATAATGTTTAAAAAGGTATT	1226
601	QY	ATTGCTTTGGGTGATGATGAACATCTACGTTTAAATTTGACGACAGATGTTCCAAATTTATTAT	660
1227	DB	ATTGCTTTGGGTGATGATGAACATCTACGTTTAAATTTGACGACAGATGTTCCAAATTTATTAT	1286
661	QY	TATGGATTTTAAAGATTCGGATGACATTTATGCTCAAAAATATTCAAATTTACGGATAAAGGT	720
1287	DB	TATGGATTTTAAAGATTCGGATGACATTTATGCTCAAAAATATTCAAATTTACGGATAAAGGT	1346
721	QY	ACTGCTTTTGATGTPATGFGAATGGTAGTTTTATGATCATCTTCCTGTCCTCCAAATAT	780
1347	DB	ACTGCTTTTGATGTPATGFGAATGGTAGTTTTATGATCATCTTCCTGTCCTCCAAATAT	1406
781	QY	GGTGACCATACAGTTTTTAAATGCAATTAGCTGTAATTTGCGATTTAGTTATTTAGAGAAGCTA	840
1407	DB	GGTGACCATACAGTTTTTAAATGCAATTAGCTGTAATTTGCGATTTAGTTATTTAGAGAAGCTA	1466
841	QY	GATGTTTACAAATATTAAAGAAAGCATTTAGAAAACGTTTTGGTGGTGTAAACCGTCTTTCAAT	900
1467	DB	GATGTTTACAAATATTAAAGAAAGCATTTAGAAAACGTTTTGGTGGTGTAAACCGTCTTTCAAT	1526
901	QY	GAACTTACAATTCGCAATCAAAGTTATTGTAGATGATTTATGCACACCATCCACAGAGAAAT	960
1527	DB	GAACTTACAATTCGCAATCAAAGTTATTGTAGATGATTTATGCACACCATCCACAGAGAAAT	1586
961	QY	AGTGCTTACAATTTGAAACAGCAGAAAGAAATATCCACATAAGAAGTTGTGTGAGTATTT	1020
1587	DB	AGTGCTTACAATTTGAAACAGCAGAAAGAAATATCCACATAAGAAGTTGTGTGAGTATTT	1646
1021	QY	CAACCCACACTTTCCTCTAGAACACAGGCATTTTTTAAATGAAATTCGAGAAAGTTTAAAGT	1080
1647	DB	CAACCCACACTTTCCTCTAGAACACAGGCATTTTTTAAATGAAATTCGAGAAAGTTTAAAGT	1706
1081	QY	AAAGCAGATCGTGATTTCTTATGTAATTTTTTGGATCAATTAGAGAAAATATCTGGCGCA	1140
1707	DB	AAAGCAGATCGTGATTTCTTATGTAATTTTTTGGATCAATTAGAGAAAATATCTGGCGCA	1766
1141	QY	TTAAACGATACAGATTTAATTTGATAAAATTTGAAGGTGCATCGTTTAATTAATGAAGATTCT	1200
1767	DB	TTAAACGATACAGATTTAATTTGATAAAATTTGAAGGTGCATCGTTTAATTAATGAAGATTCT	1826
1201	QY	ATTAAATGATTTAGAACAAATTTGATAATGCTGTTATTTTATTATGGTGAGGTGATATT	1260
1827	DB	ATTAAATGATTTAGAACAAATTTGATAATGCTGTTATTTTATTATGGTGAGGTGATATT	1886
1261	QY	CAAAAAATTACAAATGCAATTTTAGATAAATTTAGGCATGAAAAATCGTTTTTAA	1314
1887	DB	CAAAAAATTACAAATGCAATTTTAGATAAATTTAGGCATGAAAAATCGTTTTTAA	1940
RESULT 3			
AX742071			

Db	582	TATGGATTTTAAAGATTCCGATGCATTTATATGCTCAAAATATTCAAATTTACGGATAAAGGT	741
Qy	721	ACTGCTTTTGGATGTGTGATGGTGTGATTTTATGATCATCTTCTCTCTCCCAAAAT	780
Db	742	ACTGCTTTTGAATGTGTGATGGTGTGATTTTATGATCATCTTCTCTCTCCCAAAAT	801
Qy	781	GGTGACCATACAGTTTAAATGCATTAGCTGTAAATGGCATTTAGTATTATTAGAGAAGCTA	840
Db	802	GGTGACCATACAGTTTAAATGCATTAGCTGTAAATGGCATTTAGTATTATTAGAGAAGCTA	861
Qy	841	GATGTTTACAAATATTAAAGAAGCATTTAGAAAAGTTTGGTGGTGTAAAGCTGTTTCAAT	900
Db	862	GATGTTTACAAATATTAAAGAAGCATTTAGAAAAGTTTGGTGGTGTAAAGCTGTTTCAAT	921
Qy	901	GAACCTACAAATTCGAATCAAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAAT	960
Db	922	GAACCTACAAATTCGAATCAAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAAT	981
Qy	961	AGTGCTACAAATTTGAAAACAGCACGAAAAGAAATATCCACATAAAGAAGTTGTTGCAATTTT	1020
Db	982	AGTGCTACAAATTTGAAAACAGCACGAAAAGAAATATCCACATAAAGAAGTTGTTGCAATTTT	1041
Qy	1021	CAACCCACACTTTCTCTAGAAACACAGGCATTTTTTAAATGAATTTGCAGAAAGTTTAAGT	1080
Db	1042	CAACCCACACTTTCTCTAGAAACACAGGCATTTTTTAAATGAATTTGCAGAAAGTTTAAGT	1101
Qy	1081	AAAGCAGATCGTGATTTCTTTATGATGAAAATTTTTTGGATCAATTTAGAGAAAATATCTGCGCA	1140
Db	1102	AAAGCAGATCGTGATTTCTTTATGATGAAAATTTTTTGGATCAATTTAGAGAAAATATCTGCGCA	1161
Qy	1141	TTAAACGATACAAGATTTAAATGTGATAAAATTTGAAGGTGCATCGTTAATTAATGAAGATTC	1200
Db	1162	TTAAACGATACAAGATTTAAATGTGATAAAATTTGAAGGTGCATCGTTAATTAATGAAGATTC	1221
Qy	1201	ATTAAATGTAATTAGAACAAATTTGATAATGCTGTTATTTTATTATGGGTGCAGGTGATTT	1260
Db	1222	ATTAAATGTAATTAGAACAAATTTGATAATGCTGTTATTTTATTATGGGTGCAGGTGATTT	1281
Qy	1261	CAAAAATTACAAAATGCATATTTAGATAAAATTAGGCATGAAAAATCGTTTTTAA	1314
Db	1282	CAAAAATTACAAAATGCATATTTAGATAAAATTAGGCATGAAAAATCGTTTTTAA	1335

```

RESULT 4
AX617936
LOCUS      1332 bp      DNA      linear
DEFINITION      Sequence 899 from Patent WO02094868.
ACCESSION      AX617936
VERSION        AX617936.1  GI:28448287
KEYWORDS
SOURCE
ORGANISM      Staphylococcus aureus
              Staphylococcus aureus
              Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE      1
AUTHORS      Masignani, V.C., Mora, M.C. and Scarselli, M.C.
TITLE        Staphylococcus aureus proteins and nucleic acids
JOURNAL      Patent: WO 02094868-A 699 28-NOV-2002;
              Chiron Spa (IT)
FEATURES
            source
            1..1332
            /organism="Staphylococcus aureus"
            /mol_type="unassigned DNA"
            /db_xref="taxon:1280"
ORIGIN
Query Match      99.5%;      Score 1311;      DB 6;      Length 1332;
Best Local Similarity 100.0%;      Pred. No. 1.6e-189;
Matches 1311;      Conservative 0;      Mismatches 0;      Indels 0;      Gaps 0;

1  ATGCACACATATCATTTGTCGGAAATTAAAGTTCTCGCATGAGTTCATTAGCACAATC 60
22 ATGCACACATATCATTTGTCGGAAATTAAAGTTCTCGCATGAGTTCATTAGCACAATC 81

```

QY	61	ATGCATGATTTAGGACATCAAGTCTCAAGATCGGATATTGAGAACTACGTATTTACAGAA	120
DB	82	ATGCATGATTTAGGACATCAAGTCTCAAGATCGGATATTGAGAACTACGTATTTACAGAA	141
QY	121	GTTGCTCTTAGAATAAAGGGGATAAAATATTTACCATTTTGATGCTAAATAACAATAAAGAA	180
DB	142	GTTGCTCTTAGAATAAAGGGGATAAAATATTTACCATTTTGATGCTAAATAACAATAAAGAA	201
QY	181	GATATCGTAGTTATATACAAGGTAAATGCAATTCGGAGTAGCCATGAAGAAATAGTAGCTGA	240
DB	202	GATATCGTAGTTATATACAAGGTAAATGCAATTCGGAGTAGCCATGAAGAAATAGTAGCTGA	261
QY	241	CATCAATGAAATGATAGTGTGTAAGTATATATGATTTTTTATAGGACAGATTAATGATCAA	300
DB	262	CATCAATGAAATGATAGTGTGTAAGTATATATGATTTTTTATAGGACAGATTAATGATCAA	321
QY	301	TATACCTTCAGTAGCTGTAACTGCTGCACATGGTAAAACTTCTACAACAGGTTTATTTATCA	360
DB	322	TATACCTTCAGTAGCTGTAACTGCTGCACATGGTAAAACTTCTACAACAGGTTTATTTATCA	381
QY	361	CATGTTATGAATGGTGATMAAAAGACTTCATTTTTTAATTTGGTGATGGCAAGGTATGGGA	420
DB	382	CATGTTATGAATGGTGATMAAAAGACTTCATTTTTTAATTTGGTGATGGCAAGGTATGGGA	441
QY	421	TTGCCTGAAAGTGATTAATTTTCGCTTTTGGAGCATGTGAATATAGAGCTCACATTTTAACT	480
DB	442	TTGCCTGAAAGTGATTAATTTTCGCTTTTGGAGCATGTGAATATAGAGCTCACATTTTAACT	501
QY	481	TATAAACCTGATTCAGCAATATATGACAAATATTGATTTTCGATCATCTGATTAATTTTAAA	540
DB	502	TATAAACCTGATTCAGCAATATATGACAAATATTGATTTTCGATCATCTGATTAATTTTAAA	561
QY	541	GATATTAATGATGTTTTTGATGCAATTCCAAGAAATGGCACATAAATGTTAAAAAGGTATT	600
DB	562	GATATTAATGATGTTTTTGATGCAATTCCAAGAAATGGCACATAAATGTTAAAAAGGTATT	621
QY	601	ATTGCTTTGGGTGATGATGAACATCTACGTAAATTTGAACAGAGTGTCCAAATTTATTAT	660
DB	622	ATTGCTTTGGGTGATGATGAACATCTACGTAAATTTGAACAGAGTGTCCAAATTTATTAT	681
QY	661	TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTCAGGATAAAGGT	720
DB	682	TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTCAGGATAAAGGT	741
QY	721	ACTGCTTTTGATGTGATGTGGATGGTGAAGTTTTATGATCACTTCCTGTCCACAATAT	780
DB	742	ACTGCTTTTGATGTGATGTGGATGGTGAAGTTTTATGATCACTTCCTGTCCACAATAT	801
QY	781	GGTGACCAATACAGTTTTTAATGCAATTAGCTGTAAATTCGGATTAGTATTTAGAGAGCTA	840
DB	802	GGTGACCAATACAGTTTTTAATGCAATTAGCTGTAAATTCGGATTAGTATTTAGAGAGCTA	861
QY	841	GATGTTACAAATATTTAAAGAGCAATAGAAAGTTGCTGTGTGTTAAACGTCGTTTCAAT	900
DB	862	GATGTTACAAATATTTAAAGAGCAATTAGAAAGTTGTTAAAGTTGTTAAAGTTTCAAT	921
QY	901	GAAACTACAATTTGCAAAATCAAGTTATTTGTAGATGATTTATGCACACCATCCAGAGAAATT	960
DB	922	GAAACTACAATTTGCAAAATCAAGTTATTTGTAGATGATTTATGCACACCATCCAGAGAAATT	981
QY	961	AGTGCTACAAATGTAACAGACGAAAGAAATATCCACATAAAGAGTGTGTCAGTATTT	1020
DB	982	AGTGCTACAAATGTAACAGACGAAAGAAATATCCACATAAAGAGTGTGTCAGTATTT	1041
QY	1021	CAACACACACTTTCTCTAGAACACAGGCATTTTTTAAATGAATTTGACAGAAAGTTTAACT	1080
DB	1042	CAACACACACTTTCTCTAGAACACAGGCATTTTTTAAATGAATTTGACAGAAAGTTTAACT	1101
QY	1081	AAAGCAGATCGTGATTTCTTATGTGAAATTTTGGATCAATTTAGAGAAATATCTGGCGCA	1140
DB	1102	AAAGCAGATCGTGATTTCTTATGTGAAATTTTGGATCAATTTAGAGAAATATCTGGCGCA	1161
QY	1141	TTAACGATACAAGATTTTAATTTGATATAAAATGGAAGGTGCATCGTTAATTAATGAAGATCT	1200

```
Db 1162 TTAACGATCAAGATTAAATGATATAAATGAAGGTCATCGTTAATTAATGAAGATTCT 1221
Qy 1201 ATTAATGATTAGAACAAATTTGATATGCTGTTATTTATTTATGCGTGCAGGTGATATT 1260
Db 1222 ATTAATGATTAGAACAAATTTGATATGCTGTTATTTATTTATGCGTGCAGGTGATATT 1281
Qy 1261 CAAAAATTACAAAATGCATATTAGATAAATAGGATGAAAATGCGTTT 1311
Db 1282 CAAAAATTACAAAATGCATATTAGATAAATAGGATGAAAATGCGTTT 1332

RESULT 5
AP003134/2 301550 bp DNA linear BCT 24-APR-2003
LOCUS Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
DEFINITION genome, section 6/10.
ACCESSION AP003134 BA000018
VERSION AP003134.2 GI:14349226
KEYWORDS
SOURCE Staphylococcus aureus subsp. aureus N315
ORGANISM Staphylococcus aureus subsp. aureus N315
REFERENCE 1
AUTHORS Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I.,
Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M.,
Matsumaru, H., Maruyama, A., Murakami, H., Hosoyama, A.,
Mizutani, U., Takahashi, N. K., Sawano, T., Inoue, R., Kaito, C.,
Sekimizu, K., Hirakawa, H., Kihara, S., Goto, S., Yabuzaki, J.,
Kanehisa, M., Yanashita, A., Oshima, K., Furuya, K., Yoshino, C.,
Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. and Hiramatsu, K.
TITLE Whole genome sequencing of methicillin-resistant Staphylococcus
aureus
JOURNAL Lancet 357 (9264), 1225-1240 (2001)
MEDLINE 21311952
PUBMED 11418146
REFERENCE 2 (bases 1 to 301550)
AUTHORS Hosoyama, A., Nagai, Y., Kuroda, M., Hiramatsu, K. and Oguchi, H.
TITLE Direct Submission
JOURNAL Submitted (30-JAN-2001) Director-General, Biotechnology Center,
National Institute of Technology and Evaluation, Biotechnology
Center; 4-1-8 Higashi, Shiga, 516-8580, Japan
COMMENT (E-mail: bio.nite.go.jp; URL: http://www.bio.nite.go.jp/)
Tel: 81-3-3481-1933 Fax: 81-3-3481-8424
On Jun 12, 2001 this sequence version replaced gi:13701258.
FEATURES
Location/Qualifiers
1..301550
/organism="Staphylococcus aureus subsp. aureus N315"
/mol_type="genomic DNA"
/strain="N315"
/sub_species="aureus N315"
/db_xref="taxon:158879"
complement (249..947)
/gene="SA1293"
complement (249..947)
/gene="SA1293"
/feature="ORFID:SA1293"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB42553.1"
/db_xref="GI:13701259"
/translation="MSEFTNIIYFVLMVPLVLAQHKVSNVEKYSQVSTSGTKGR
EVALEILHANGIYDVEVKGEGFLTDHYDPKKVSLSPANRPSVAGTAAIAHEVG
HAIQDQGYWFLFRALVPMVNLGSLSYIMILGILTAIGSFGSTALMGAGLM
SLAVLSIVTLPEVFDASSRAMQITALNIVNEKEYKHARKVLSAAATVATAV
AELVRIILARSDD"
complement (1001..1588)
/gene="SA1294"
complement (1001..1588)
/gene="SA1294"
/feature="ORFID:SA1294"
```

```
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB42554.1"
/db_xref="GI:13701260"
/translation="MTINTFWQYTYIYQSRWMLMLLNCILMIYGIWYGEOLSHTPW
HFKIFPDSPTAILYVISILILIQKNSIIEALAFVTLFKYGIWAVIMNLFLEIQ
GDITVNGLVLMFHSINAVQAIYVPRKSRMIGISVAMIVFLNDIYDIYFHLQFPY
DFITVWQIGVLSCLSVFGLLYIELNKLKCK"
complement (1578..2153)
/gene="SA1295"
complement (1578..2153)
/gene="SA1295"
/feature="ORFID:SA1295"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB42555.1"
/db_xref="GI:13701261"
/translation="MSETLQIKESFIEYLLFOYRFSKRIAVWLVNLYKNEAKLANI
HFVDTKINHETLEIAVGVSHASAIQTKNIKMLNTNEIDYIANTHCAFDIQHFA
NVSKRQRLDLIVLAQTLSPSYQTYLHLNLSMAIDRRHKALLIDYLLHNDLSLQMN
EKORFYQLTQILNTLKLNVKNQFEDLADD"
complement (2167..3411)
/gene="SA1296"
complement (2167..3411)
/gene="SA1296"
/feature="ORFID:SA1296"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB42556.1"
/db_xref="GI:13701262"
/translation="MEDIYKLIIDDINLQKLENDLSRVNEAITDNDALFILGETLYN
FGMPQLEVFRVLYHKYPDESELLYFIEGLMSENOTDEALEYLVSPSPKMLE
ADLYQINMVEAIDKQALEPNDPIIHFALAEMLYDYDQVLRATSEVETWLC
EYQVNGWLFMRMADCSLOSGNYSDAIRLYDEINEDMTSEDYLLKASIDKNDITQE
AKIMTLLSKDPDYIQGHYQSLYENENYYPDAIETGKGLKLSQFYKELMYTTC
LEIHGDANEVGLLQKALEVNAVYQEPPLISLYRNEEDYEAIIELLYVDEEDLD
PTFMHFAFYQOERDKQAFPELAYPTMTNIDFMSDYFYFLIEIGQKEQAITIL
NOLLEPSENWHDLSRLQS"
complement (3418..4716)
/gene="aroA"
complement (3418..4716)
/gene="aroA"
/feature="ORFID:SA1297"
/codon_start=1
/transl_table=11
/product="3-phosphohikimate 1-carboxyvinyltransferase"
/protein_id="BAB42557.1"
/db_xref="GI:13701263"
/translation="NVSEQIIDIISGLKGEIEVPGDKSMTHRAIMLASLAEGTSNIYK
PLLGDCRTWMDIFRLGVDIKEDDKLVNPSGYKAKTPOHVLTYGSGTTRLLIA
LLGSLGTSVLSGDVSTGKRPMDVLRPLKMDANIEGIEDNYTLLIKPSVIKGIN
YQMEVASQVKSAILPASLFNDTIVKELDVSRNHTETMFRHNPIDKRSUTTT
PDAIOHIAFPDHPVGDISSAFAFFVAALITPESDVTHNVGINPFRSGIDIVKMG
GNIQLFNQTTGAPTASIRIQYTPMLQPTITIEGELVPKALDELPLVLLCTQAVGST
IKDAELKVKETNRIDTADMLNLGLFELQPTNDGLIHPSEKTNATVDSLTDHRIIG
MMLAVASLSSPEPVKIQFDVAVNVSPFGPLPKLLENEG"
complement (4726..5790)
/gene="aroB"
complement (4726..5790)
/gene="aroB"
/feature="ORFID:SA1298"
/codon_start=1
/transl_table=11
/product="3-dehydroquinase"
/protein_id="BAB42558.1"
/db_xref="GI:13701264"
/translation="MKLQTYPSNNYPIFVEHGAIDHISTYIDQFDSFILIDEHVNO
YFADKFDLISYENVHVKVILIPAGEKTKFEQYQETLEILSHVTRNTAIIAUGGAT
GDFAGFAATLIRGVHFIQVPTTILAHDSVGVKINSQKGNLIGAIYRPTAVIYD
LDLFLKTLFPEQILSGYAEVYKHALINGESTTQETIEBQHKDRILQSLNGMDKYIANGI
```

```
ETKLDIVVDEKQCVKFLNLGHTFHGAVENYHKAIGHAVMTGIIVQIVANILFN
SNHDIQYINILKLGPLETITDIDPETIYQYMLSKNDKQGVQWVLLIKHFGDIVV
QHIDQTLQACQKQKLYFK"
complement (5816. .6982)
/gene="aroc"
complement (5816. .6982)
/gene="aroc"
/note="ORFID:SA1299"
/codon_start=1
/transl_table=11
/product="chorismate synthase"
/protein_id="BAB42559.1"
/db_xref="GI:13701265"
/translation="MRILTSGSEHPQLTWIVGIPANLEIKVEDINKMPKRGQYV
RGRMQLEKDTIVSVGRNGYTLGSPITWVTVNDFTWIKMGAAPISBEERENMK
RTTKPRFGHADIVLGMKYNLDRNLVRSARETAARVAVAGVLLQOLDIY
SRVVEIGIKDQFYDSETPKANLDRNDVRIDDSIAQMEDKIDEAKNEDSIGGV
QVVENMPYGVSVHVDKLDCKIAQGVVINA PKGVSPGEGKAAKPGSEIQDEI
LYNSEIGYRGSHLGGLEGMNGMFLIIVGVNPKPIPTLYKPLNSVDINTKDFKAT
IERSDCAVPAASIVCHVVAFEIAKALLEFQSNHIEQLQQQIADRRQLNVF"
complement (7368. .7568)
/gene="SA1300"
complement (7368. .7568)
/gene="SA1300"
/note="ORFID:SA1300"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB42560.1"
/db_xref="GI:13701266"
/translation="MKNISIIILGILMFLKMLBLLYGAIFLDKPLNPITKIIFILTLI
YIFVLVKSLLIIFLSKYNKSA"
complement (7597. .7722)
/gene="SAS045"
complement (7597. .7722)
/gene="SAS045"
/note="ORFID:SA045"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB42561.1"
/db_xref="GI:13701267"
/translation="MWLVIMIHNKARFLCCSKLEKFKYFLVTRINKRGIFFIK"
complement (7777. .8226)
/gene="ndk"
complement (7777. .8226)
/gene="ndk"
/note="ORFID:SA1301"
/codon_start=1
/transl_table=11
/product="nucleoside diphosphate kinase"
/protein_id="BAB42562.1"
/db_xref="GI:13701268"
/translation="MERTFMILKPDVQRNLIGEVISRIERKGLKLVGKLMQVPMEL
AETHYGEHQKPFIDLSIFTSAPVPMVVEGEDAVNYSRIHIGSNPSEAPSGSIR
GDLGLTVGNRIIHGSDLSASREINLWENENEITSYASPRDAWLYE"
complement (8318. .9277)
/gene="gerCC"
complement (8318. .9277)
/gene="gerCC"
/note="ORFID:SA1302"
/codon_start=1
/transl_table=11
/product="heptaprenyl diphosphate synthase component II"
/protein_id="BAB42563.1"
/db_xref="GI:13701269"
/translation="MAKLMMNEIKKVEORLEKAKSKDSVLEQASLHLLSSGKRVK
PAFVILSQFGKDEQTEQTYQVAVALELIHMAVLVHDDVIDKSKRERKLIISKWD
```

Query Match

Best Local Similarity 99.5%; Score 1310.8; DB 1; Length 301550;

Matches 1312; Conservativity 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGACACACTATCATTTTCTCGGAATTAAGGTTCTGGCATGAGTTCATTAGACAAATC 60
Db 291118 ATGACACACTATCATTTTCTCGGAATTAAGGTTCTGGCATGAGTTCATTAGACAAATC 291059
QY 61 ATGCAATGATTAGGACATGAAGTTCAAGGATCGGATATGAGAACTACGTTATTTACAGAA 120
Db 291058 ATGCAATGATTAGGACATGAAGTTCAAGGATCGGATATGAGAACTACGTTATTTACAGAA 290999
QY 121 GTTGCTCTTAGAAATAAGGGGATAAAAAATATACCAATTTGATGCTAATACATAAAGAA 180
Db 290998 GTTGCTCTTAGAAATAAGGGGATAAAAAATATACCAATTTGATGCTAATACATAAAGAA 290939
QY 181 GATATGGTAGTTTATACAAGTAAATGCAATCGGAGTAGCCATCAAGAAATAGTACGTGCA 240
Db 290938 GATATGGTAGTTTATACAAGTAAATGCAATCGGAGTAGCCATCAAGAAATAGTACGTGCA 290879
QY 241 CATCAATTGAATTAGATGTTGTAAGTTAATAAGTATTTTATAGGACAGATTTATGATCAA 300
Db 290878 CATCAATTGAATTAGATGTTGTAAGTTAATAAGTATTTTATAGGACAGATTTATGATCAA 290819
QY 301 TATACCTTCAGTAGCTGTAACCTGGTGCACATGCTTAAACTCTTACACACAGTTTATTTATCA 360
Db 290818 TATACCTTCAGTAGCTGTAACCTGGTGCACATGCTTAAACTCTTACACACAGTTTATTTATCA 290759
QY 361 CATGTTATCAATGGTGATAAAAAAGACTTTCATTTTAAATTTGGTGATGGCAAGGTATGGGA 420
Db 290758 CATGTTATCAATGGTGATAAAAAAGACTTTCATTTTAAATTTGGTGATGGCAAGGTATGGGA 290699
QY 421 TTGCTCTGAAAGTATTTTTCGCTTTGAGGCATGTGAATATAGACGTCACCTTTTAAAGT 480
Db 290698 TTGCTCTGAAAGTATTTTTCGCTTTGAGGCATGTGAATATAGACGTCACCTTTTAAAGT 290639
QY 481 TATAAACCTGATACGCAATTTATGACAAATTTGATTTTCGATCATCTCTGATTTTATAA 540
Db 290638 TATAAACCTGATACGCAATTTATGACAAATTTGATTTTCGATCATCTCTGATTTTATAA 290579
QY 541 GATATTAATGATGTTTTTTGATGCAATCCAGAAATAGGCACATAAATGTTAAAAAGGTATT 600
Db 290578 GATATTAATGATGTTTTTTGATGCAATCCAGAAATAGGCACATAAATGTTAAAAAGGTATT 290519
QY 601 ATTGCTTTGGGTGATGATGAACATCTACGTAAAAATGAAGCAGATGTTCCAAATTTATTAT 660
Db 290518 ATTGCTTTGGGTGATGATGAACATCTACGTAAAAATGAAGCAGATGTTCCAAATTTATTAT 290459
QY 661 TATGGAATTAAGATTCGGATGACATTTATGCTCAAATATTTCAAATTCAGGATAAAGT 720
Db 290458 TATGGAATTAAGATTCGGATGACATTTATGCTCAAATATTTCAAATTCAGGATAAAGT 290399
QY 721 ACTGCTTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 290398 ACTGCTTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 290339
QY 781 GGTGACCATACAGTTTTTAAATGCAATTTAGCTGTAATTTGCGATTAGTTATTTAGAGAAGCTA 840
Db 290338 GGTGACCATACAGTTTTTAAATGCAATTTAGCTGTAATTTGCGATTAGTTATTTAGAGAAGCTA 290279
QY 841 GATGTTCAAATATTAAGAAAGCATTAGAAACGTTGGTGGTGTGTTAAACGTCGTTTCAAT 900
Db 290278 GATGTTCAAATATTAAGAAAGCATTAGAAACGTTGGTGGTGTGTTAAACGTCGTTTCAAT 290219
QY 901 GAAACTACAAATTCGAAATCAAGTTATTTGATGATGATGATGATGATGATGATGATGATGAT 960
Db 290218 GAAACTACAAATTCGAAATCAAGTTATTTGATGATGATGATGATGATGATGATGATGATGATGAT 290159
QY 961 AGTGCTACAAATTCGAAACAGCAGCAAGAAATATCCACATAAAGAAAGTTGTTGCGAGTATTT 1020
Db 290158 AGTGCTACAAATTCGAAACAGCAGCAAGAAATATCCACATAAAGAAAGTTGTTGCGAGTATTT 290099
QY 1021 CAACCAACACATTTTCTCTAGAACACAGGCAATTTTAAATGAATTTGCAGAAAGTTTAAAGT 1080
Db 290098 CAACCAACACATTTTCTCTAGAACACAGGCAATTTTAAATGAATTTGCAGAAAGTTTAAAGT 290039
QY 1081 AAAGCAGATCGTGTATCTTATGTGAAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1140
```



```
gene
CDS
KIYEYENDALRAYGLTVKTEGFGTHMNVSNNDGPVTIIYESDQGIQ"
complement (6118..8307)
/gene="relA"
complement (6118..8307)
/gene="relA"
/note="SAV1634"
/codon_start=1
/transl_table=11
/product="GTP pyrophosphokinase"
/protein_id="BAB57796.1"
/db_xref="GI:14247406"
/translation="MNNEYPYSADEVILHKAKSYSLSADEYEVILKSVHIAVEAHKGOPR
KNGLPVIMHPIDVAGLITSEMRLDPTIVAGFLHDVIEDPYTSEDVKEMENRVARIV
DGYTKLKYKYSRKEQOENRHLKFIATAKQVRIIVLKLADLHNRILKAMPREKQ
IKRISRTDEIYLAHLHGINTIKWELEDALRIIDNVQIFRIIVNLMLKRSREAYI
DCAILDIRTEMNMTESDINGRPHIYIYRKMKQKQFQDQIFDLAIRIVINSIN
EYAILGLVHTTILKMPGFEKQIAMPKQNLQYSLHTTVVPGNDPLEIQIRTFDMEH
IAEHGVAHWAHYKEGKVEKQDQYQKLNWLKELAEADHTSDAQBEMETLAKYDQS
DKVYATPASDVIELPYGAVPIDPAVAIHSEGVNKMIGAKVNGKIPIPIDYILQTDIV
ERTSKHSGPRDWLKIYKSSAKKIISFFPKQDRSSNIEKGRMMVEVEIKQGFPR
VEDIUTERNKIYVNEKINFANEDDLFAAVFGFVTSLOIVNKLTERQRIILDKQRLNE
EGEVTSLEPIKONIIDSQVYVEGLENVLIKLSCCNPIPGDDIVGYITKGHGIKYHR
TDCPNIKNETERLINVWYKSDATQYQVDLEVTAYDRNGLLNEVLQAVSSTAGNLI
KYSGRSDIDKNAINISVMVKNVNDVYRVEKIKQLGDVYTVTVRVN"
complement (8735..9253)
/gene="apt"
complement (8735..9253)
/gene="apt"
/note="SAV1635"
/codon_start=1
/transl_table=11
/product="adenine phosphoribosyl transferase"
/protein_id="BAB57797.1"
/db_xref="GI:14247407"
/translation="MDLKQYSEVQDMPKPGVSFKDITTIMDNGEAYGVATDKIVEYA
KRDYDIVVGPARGFIICPWAYSNGIIGFAPVRKEGKLPREVIRYEYDLEYKTIVLT
MHKDALKQQRVLITDDLLATGCTIEAAIKLVEKLGIVVGIAFIILKYLINGIEKIK
DYDWSLISYDE"
complement (9275..11548)
/gene="SAV1636"
complement (9275..11548)
/gene="SAV1636"
/note="similar to single-strand DNA-specific exonuclease"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57798.1"
/db_xref="GI:14247408"
/translation="MKPKYKWLTKPAEYISDELTSKULKTPIVKILLESKSIIDBQ
ALIESIISDQVNDHALQSLDMTKAIERIKRAIANDBEKILVYDGDVTDVSTILVTT
LQLLQAGQVHNPRTFEGYGNELAFRNAHDEGITLIITVDNGIQGHNEIKRVODLG
VDVIVTDHIEIGSTPEAYAIYHPMHPSPNYFPQQLCGAGVAKLAQALINVPDYFK
ALVAIGTTADLVSLTDENSLVKQGLKLVNDQOPTSVKALLKEAGVNDNIDETIGFI
IGPRNAVGRLODASLACELMTDDEEAAFLAEQVEHFNREKDIIVANITEAMAMA
ETRVKKGDLFILLAKHLLWGDVIGVAKIVETFAUPLILNDRQNHKAGSRSID
QVSMPEILSAHQELIAKFGHMAAGMTMDIENIESLAEGLNKWMLSKTISLSDPVK
PVDVLLTENDITIKNIRDMNRLEPFGTDFSRPIFEMDDLSVSVKAIQOQKNHLKLT
GESNTAALFWQNGHPELQDEQPINILGSVQINENWNGOSPOLIIOIDIAENQQILID
YRSKRSLPFTENDENIVLIIHPKSDKVNANVYGEIKQQDQKVLRLDPLTSMEDL
SNSLQQLQPSQLYIVLQHNHSIYFDGIPNWDIPFKYKALITIKQETNIQKGMLLCQH
LSVKPDTLKFMLKVLFDLKFVTEQDGLIRINQOQPKRSIDSSKVVYQLRQOQRMDEVKQL
LYQDFSEIKNWKSKQS"
complement (11751..14030)
/gene="secf"
complement (11751..14030)
/gene="secf"
/note="SAV1637"
/codon_start=1
/transl_table=11
/product="protein-export membrane protein SecDF"
/protein_id="BAB57799.1"
/db_xref="GI:14247409"
```

```
Query Match 99.5%; Score 1310.8; DB 1; Length 342600;
Best Local Similarity 99.8%; Pred. No. 3.3e-190;
Matches 1312; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCACACTATCATCTTTGTCGGAATAAAGTTCCTGCGATGAGTTTCATTAGCACAAATC 60
DB 132799 ATGCACACTATCATCTTTGTCGGAATAAAGTTCCTGCGATGAGTTTCATTAGCACAAATC 132740

QY 61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTTGAGAACTACGTAATTTACAGAA 120
DB 132739 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTTGAGAACTACGTAATTTACAGAA 132680

QY 121 GTTGCTCTTAGAATAAGGGATAAAAATATACATTTGATGCTTAATAACAATAAAGAA 180
DB 132679 GTTGCTCTTAGAATAAGGGATAAAAATATACATTTGATGCTTAATAACAATAAAGAA 132620

QY 181 GATATGGTAGTTATACAAAGGTAATGCTTCGCCAGTAGCCATGAAGAATACTAGCTGCA 240
DB 132619 GATATGGTAGTTATACAAAGGTAATGCTTCGCCAGTAGCCATGAAGAATACTAGCTGCA 132560

QY 241 CATCAATTTGAAATTAGATGTTTGAAGTTATATGATTTTTTAGGACAGATTTATGATCAA 300
DB 132559 CATCAATTTGAAATTAGATGTTTGAAGTTATATGATTTTTTAGGACAGATTTATGATCAA 132500

QY 301 TATACCTTCAGTAGCTGTAACTGGTGCACTGGTAAAACTCTTACAAACAGGTTTATATCA 360
DB 132499 TATACCTTCAGTAGCTGTAACTGGTGCACTGGTAAAACTCTTACAAACAGGTTTATATCA 132440

QY 361 CATGTTATGAATGCTGTATAAAAAGACTTCATTTTAAATTTGGTGAGGCACAGATATGGA 420
DB 132439 CATGTTATGAATGCTGTATAAAAAGACTTCATTTTAAATTTGGTGAGGCACAGATATGGA 132380

QY 421 TTGCTCTGAAAGTAGTATTTTCGCTTTTGAGGCATGCTGAATATAGACGTCACATTTTAACT 480
DB 132379 TTGCTCTGAAAGTAGTATTTTCGCTTTTGAGGCATGCTGAATATAGACGTCACATTTTAACT 132320

QY 481 TATAAACCTGATACGCAATATATGCAAAATATGATTTTCGATCATCTCGATATTTTAAA 540
DB 132319 TATAAACCTGATACGCAATATATGCAAAATATGATTTTCGATCATCTCGATATTTTAAA 132260

QY 541 GATATTAATGATGTTTGTGATGCTCCAGAAATGGCACATAATGTTAAAAAAGGTATT 600
DB 132259 GATATTAATGATGTTTGTGATGCTCCAGAAATGGCACATAATGTTAAAAAAGGTATT 132200

QY 601 ATTCGTTGGGTGATGATGAACATCTACGTAAAAATTTGAAGCAGATGTTCCCAATTTATAT 660
DB 132199 ATTCGTTGGGTGATGATGAACATCTACGTAAAAATTTGAAGCAGATGTTCCCAATTTATAT 132140

QY 661 TATGGATTTAAAGATTCCGATGACATTTATGCTCAAAATATTTCAAATTCGGATAAAGT 720
DB 132139 TATGGATTTAAAGATTCCGATGACATTTATGCTCAAAATATTTCAAATTCGGATAAAGT 132080

QY 721 ACTGCTTTGATGATGATGAGGTGAGTTTATGATCATCTCCTGCTCCCAATAT 780
DB 132079 ACTGCTTTGATGATGATGAGGTGAGTTTATGATCATCTCCTGCTCCCAATAT 132020

QY 781 GGTGACCATACAGTTTAAATGCAATAGCTGTAATTTGCGATTAGTTATTTAGAGAAGCTA 840
DB 132019 GGTGACCATACAGTTTAAATGCAATAGCTGTAATTTGCGATTAGTTATTTAGAGAAGCTA 131960

QY 841 GATGTTTCAAATATTAAGAAGCAATTAGAAACGTTTGGTGGTGTAAACGTCGTTTCAAT 900
DB 131959 GATGTTTCAAATATTAAGAAGCAATTAGAAACGTTTGGTGGTGTAAACGTCGTTTCAAT 131900

QY 901 GAAACTCAATTTGCAAAATCAAGTTTATTTAGATGATTTATGACACCATCCAGAGAAATTT 960
DB 131899 GAAACTCAATTTGCAAAATCAAGTTTATTTAGATGATTTATGACACCATCCAGAGAAATTT 131840

QY 961 AGTGCTCAATTTGAAACAGCACGAAAGAAATATCCACATTAAGAGTTGTTGTCAGTATTT 1020
DB 131839 AGTGCTCAATTTGAAACAGCACGAAAGAAATATCCACATTAAGAGTTGTTGTCAGTATTT 131780

QY 1021 CAACCAACACTTTCTCTAGAAACAGGCAATTTTAAATGAATTTGACAGAAAGTTTAAAGT 1080
```


|||||
Db 131779 CAACCAACACTTCTCTAGAACCAAGCATTTTAAATGAATTTGCAGAAAGTTTAAGT 131720
QY 1081 AAAGCAGATCGTGTATCTTATGTAATTTTTCGATCAATTTAGAGAAATACCTGGCCCA 1140
Db 131719 AAAGCAGATCGTGTATCTTATGTAATTTTTCGATCAATTTAGAGAAATACCTGGCCCA 131660
QY 1141 TTAACGATCAAGATTTAAATTTAGTAATAATGAAGTGCATCGTTAAATTAATGAAGATCTT 1200
Db 131659 TTAACGATCAAGATTTAAATTTAGTAATAATGAAGTGCATCGTTAAATTAATGAAGATCTT 131600
QY 1201 ATTAATGATTTAGAACATTTGATATGCTGTTATTTATTTATGAGTGCAGGTGATATT 1260
Db 131599 ATTAATGATTTAGAACATTTGATATGCTGTTATTTATTTATGAGTGCAGGTGATATT 131540
QY 1261 CAAAAATTACAAATGATATTTAGATAAATTTAGGCATGAAGAAATGCGTTTAA 1314
Db 131539 CAAAAATTACAAATGATATTTAGATAAATTTAGGCATGAAGAAATGCGTTTAA 131486

RESULT 7
AP004828/c
LOCUS 272850 bp DNA linear BCT 20-DEC-2002
DEFINITION Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,
strain:MW2, section 7/10.
ACCESSION AP004828 BA000033
VERSION AP004828.1 GI:21204850
KEYWORDS
SOURCE Staphylococcus aureus subsp. aureus MW2
ORGANISM Staphylococcus aureus subsp. aureus MW2
Bacteria; Firmicutes; Bacillales; Staphylococcus.

REFERENCE
1
Baba, T., Takeuchi, F., Kuroda, M., Yuzawa, H., Aoki, K., Oguchi, A.,
Nagai, Y., Iwano, M., Asano, K., Naimi, T., Kuroda, H., Cui, L.,
Yamamoto, K. and Hiramatsu, K.
Genome and virulence determinants of high virulence
community-acquired MRSA
Lancet 359 (9320), 1819-1827 (2002)
2040717
12044378
2 (bases 1 to 272850)
Director-General, Biotechnology Center, Aoki, K., Oguchi, A.,
Nagai, Y., Asano, K., Iwano, M., Baba, T., Kuroda, M., Hiramatsu, K. and
Kikuchi, H.
Direct Submission
Submitted (06-MAR-2002) Director-General, Biotechnology Center,
National Institute of Technology and Evaluation, Biotechnology
Center; 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
(E-mail:bio@nitech.go.jp, URL:http://www.bio.nitech.go.jp/
Tel:81-3-3481-1933, Fax:81-3-3481-8424)
Location/Qualifiers
1. 272850
/organism="Staphylococcus aureus subsp. aureus MW2"
/mol_type="genomic DNA"
/strain="MW2"
/db_xref="taxon:196620"
complement (469. .1833)
/gene="MW1681"
complement (469. .1833)
/genes="MW1681"
/notes="ORFID:MW1681"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB95546.1"
/db_xref="GI:21204851"
/translation="MTQQNNKKTAKNHTYQNEPLNRKDFVVFITGALVGSALGL
YFNKVIQRADLLKVEKESQPEERKTOLEETVAFKERVGLNKSNEQAALKA
QQAIIKEEVSANLISQQAQIQEAKKEAQAEATDKSAVNEESKAAALKAQQAII
KEEASANNISDTSQEAQIQEAKKEAQAEATDKSAVNEESKAAALKAQQAIIKEEAS
ANNISDTSQEAQIQEAKKEAQAEATDKSAVNEESKAAALKAQQAIIKEEAS
OKQAKLTPGKESQLEALFAEKPAVKNDLKEIPLQVTKNDVSESTVINDKQTVK
QKAEKFGVITKAEKTNNTAVDKSKGSKKTPSKRNASKASTNKTSGQKQ

HNKSSQGAQKSSSSKSTQKNQTSNKNKSKTTNAKSSNASKTPNAPKAVKANSKIEKR
TFND"
complement (1907. .2398)
/gene="MW1682"
complement (1907. .2398)
/gene="MW1682"
/note="ORFID:MW1682
hypothetical protein, similar to general stress protein
homolog"
/codon_start=1
/transl_table=11
/protein_id="BAB95547.1"
/db_xref="GI:21204852"
/translation="MDWILPIAGITAAIAFLIICIGIVAVINSVKKNLDYVAKTLDGV
EGOVQITRETTDLLHKVNRLTEDIQGVDRLSNVVDVAVKIGDVSQVTLNSSVDRVTN
STHNISQNEDKISQVQWNSVAMEIADKQNRHRYRGRSANYKANNVATDANHSYTSR
VDK"
complement (2472. .3785)
/gene="murC"
complement (2472. .3785)
/gene="murC"
/note="ORFID:MW1683"
/codon_start=1
/transl_table=11
/product="UDP-N-Acetylmuramate-alanine ligase"
/protein_id="BAB95548.1"
/db_xref="GI:21204853"
/translation="MTHYHFVGIKSGMSLAQIMHDLGHEVQSGSDIENVVTFEVALR
NKGILPFDANNIKEDMVVIOGNAPASSHEEIVRAHQLKLDVYVSNDFIGQIIDQIT
SVATGAGHTSTTGLSHVWNGDKTSLIGDTGMGLPESDYFAFEACEYRHRFR
YKPYAINTNIDFDHPDYKQINDVDFDAFQEMAHNVKGI IAWGDHVLKRLKEADYI
YVGFKDDIYAONIQTDKTADFDVYVDFYDHFHLSPOYGDHVTNALAVIATSY
LEKLDVTNKEALETFGVKREFNETTIANOVVDYAHHPREISATIEFARKYDHK
EYVAVQHTFSRTQAFLNEFAESLSKADRVLCFEISRENTGALTIIQLIDKIEG
ASLINEDSINVLEQFDNAVFLVFGAGDIQKIQAYLIDKLGKNAF"
complement (3809. .7633)
/gene="MW1684"
complement (3809. .7633)
/gene="MW1684"
/note="DNA translocase stage III sporulation prot homolog
ORFID:MW1684"
/codon_start=1
/transl_table=11
/protein_id="BAB95549.1"
/db_xref="GI:21204854"
/translation="MSWFDKLFGEEDNSNDLIHKKRRQSSQNDIHDHDSLLPQNN
DIYSRPKRFPMSVAYENENVEQSADTIDEKEQYHRDYRQKSHDSRSQKRRRR
NQTTEOYNSBORGSKISQSIKYKDSHYHTNPKPTVYSAINGIEKTHSKTHNM
YSNNTNRKADSTPDYHKSFKTSVPSAIFCTMPKPLENGRIIPVSKPSEKVESDKO
KYDKYAKTQTSQNKQLQEKQNDVSVQGTASKSSDENSVSTTKSPNYSKVDTYIK
LENIYASQIVBEIRRRERKVKQKRFKALQKRRHKNRQSDALQALDENYAKQA
ERYVGSINDSDSLTDNSTASQLHTNGIENETVSNENKQASIQNEIDNTHVDES
PYNTEVSLNQVSTTKLSDDEVTNSVTSQHSALQHNVEVDKDELQKQSLIADS
REDGATNKEYSGSOIDDAEFVYELNDTEADTTGNIENIDNTRNASMEHVADPKTEH
AVTESQNNIDKTVNDELAPRHKQDQTNLVNSLTNDVNDGHVVDSSMNETEK
HNAETENVQNEAASEQNEVEKTIENVPKQTEKYSTLSKRPVNVMTPKKRM
DRKXKRVNVPKPVQSVSPASQAAPSSPFSQESNTNAKTNMTNNVME
NNQLIHAENDYQNAQYSEKPSADSTQTEIFEESQDDNQLEDQVQDTSVSN
EVSQITESEETHPNNTSGQDNDQKDLQPSFNQNEEDANENRPTNQDQVATN
QAVQTSKPMIRKPNIKLPSVSLLEPQVIBESDEWITDKKELNDALFVFNPAVQ
DVTEGSPVTPELSVKGVSRITAJODDIKMAAANDIRIAPIPGSRVGLVFN
QNPTVNLRSIIISPSFKNASKLIVAMGYRINNEPLIMDIATPHALLAGATGSKS
VCINSILMSLLYKNHPBELRLLLIDPRKVPFLAPYNGLPVLPVITDVAAQTSKWA
VEEMERRYKLFAYHVRNITAKKAPYDERMPKIVIVIDEADLMMAPQVEQSI
RIAOKARACGILHMLVATQPSVNVITGLIKANIPRIAPFMVSSVDSRTILDSGAER
LLQGDMLYLGSGMKNKPIRVQGTFSVSDDEIDDDVDFIKQREPDYLFEEKELLKKTQT
QSQDELDDVCAPMVNEGHSLSLQRFHQIGYNAARLIIDQLEQLGYVSSANGSKPR
DVYVTEADLNKE"
complement (7654. .8250)
/gene="MW1685"
complement (7654. .8250)
/gene="MW1685"
/note="ORFID:MW1685

phenylalanyl-tRNA synthetase (beta subunit) homolog"

```
/codon_start=1
/transl_table=11
/protein_id="BAB95550.1"
/db_xref="GI:21204855"
/translation="NNLFNPKYGVGVAPLQIEPVEGELNKNKNVVEITNEGNVVG
YNIFEISDITIEERGHKLDTDLGVNFKRISSEAGDFYKLNADLSPKFVGVYETKQ
KHPDADKLSVLNVGNDTLOIYVCGAPNVEAGQKVVYAKVGVMPGSMVIKDAELRGV
ASSGMICSMKELNLPNBEKGMVNLNDSYEIQAPFE"
complement (8279..9136)
/gene="MW1686"
CDS
complement (8279..9136)
/gene="MW1686"
/note="ORFID:MW1686"
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95551.1"
/db_xref="GI:21204856"
/translation="MNTFQMRDKLKERLSHLVDVDFKFNREBEETLRIYRTDNNKGIK
LNAIVAKYDKKEKIVDEIVYVDEAIAQMAKDTLIESISSQIMPVIRATSFDKTKQ
GVPEFYDEHTATANYAVDLKSVLELIDSLTEQOIREMSLNFVRKLSNSY
TYDEKGNFYFINSNDGDSRLINTAFNLNIEACQSGEMLVAVPHQDVLIIADIRN
KRGYDMAHLTMEFFTKGVLPIITSUSFGYKQGHLEPIFILKNNKKOKRPNVLIQLEA
NRKFNKOK"
complement (9236..9547)
/gene="MW1687"
CDS
complement (9236..9547)
/gene="MW1687"
/note="ORFID:MW1687
thioredoxin homolog"
/codon_start=1
/transl_table=11
/protein_id="BAB95552.1"
/db_xref="GI:21204857"
/translation="MKQLESEQOFESLKGQATVFETAGWCPCDRCVIEPDLPELEARY
PMFDFVSDRDRFMDICIENGIMGIPSLVYKNGELLSYIGKERKSIIEQIDFLAQY
V"
complement (9612..10688)
/gene="MW1688"
CDS
complement (9612..10688)
/gene="MW1688"
/note="ORFID:MW1688
endo-1,4-beta-glucanase homolog"
/codon_start=1
/transl_table=11
/protein_id="BAB95553.1"
/db_xref="GI:21204858"
/translation="MNINKKVTIQRITQLTELHGAPGFEEVKVNYMTQMAPYVDVFI
ENRMGGFVGKSKNPNKRVMLAAHMDIGFMTNI TKNGMIQFTNLGAVANDIQG
ORLVTKNRGDKIIGVSNIPKHFRTSGSAPDKDLTDIGAQNDEVRERIDIGD
TIVPHPTQLSEHRSYSAKANDRYGCVLAIEILKLLKDIELDVLDVLYGANNVQEEVL
RGAKASAMIDPDVAFVDCSPANDVKGSPQSLSGELGKTLIRIKDKQMLIKPFVFRDY
LLKLVFAHDIERYSPGTPDGGETHKANI GIPTAVI GVCARYIHSTDSVFDIRDY
AARSLISEAICNLNDNQIETLVK"
10775..11086
/gene="MW1689"
CDS
10775..11086
/gene="MW1689"
/note="ORFID:MW1689"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB95554.1"
/db_xref="GI:21204859"
/translation="MTFKLYIPTIIIAVAIIVISITIIQYINRKNPNPKVILNEVKS
YFMYNKGSIYVPEFVHPETDKYRLVYQGGIITIKNGQNHIDFYADAYTGEVINIVE
R"
complement (11211..12053)
/gene="MW1690"
CDS
complement (11211..12053)
/gene="MW1690"
/note="ORFID:MW1690"
```

```
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
Query Match 99.3%; Score 1309.2; DB 1; Length 272850;
Best Local Similarity 99.8%; Pred. No. 6.2e-190;
Matches 1311; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGACACACTATCATCTTTTGTGCGGAATTAAGAGTTCTGGCATGAGTTCAATAGCACAATC 60
DB 3785 ATGACACACTATCATCTTTTGTGCGGAATTAAGAGTTCTGGCATGAGTTCAATAGCACAATC 3726
QY 61 ATGCATGATTAGGACATGAAGTTCAAGATCGGATATTGAGAACTAGTATTACAGAA 120
DB 3725 ATGCATGATTAGGACATGAAGTTCAAGATCGGATATTGAGAACTAGTATTACAGAA 3666
QY 121 GTTGCTCTTAGAATAAGGGGATAAAATATTTACCATTTGATGCTATAATAACATAAAAGAA 180
DB 3665 GTTGCTCTTAGAATAAGGGGATAAAATATTTACCATTTGATGCTATAATAACATAAAAGAA 3606
QY 181 GATATGCTAGTTATACAAAGGTAATGCAITTCGCGAGTAGCCATGAAGAAATAGTACGTGCA 240
DB 3605 GATATGCTAGTTATACAAAGGTAATGCAITTCGCGAGTAGCCATGAAGAAATAGTACGTGCA 3546
QY 241 CATCAATTTGAAATTAGATGTTGTAAAGTTATAATGATTTTTTTAGGACAGATTTTGATCAA 300
DB 3545 CATCAATTTGAAATTAGATGTTGTAAAGTTATAATGATTTTTTTAGGACAGATTTTGATCAA 3486
QY 301 TATACITTCAGTAGCTGTAACTGGTGCCATCGTAAACCTTCTTACCAAGGTTTATTATCA 360
DB 3485 TATACITTCAGTAGCTGTAACTGGTGCCATCGTAAACCTTCTTACCAAGGTTTATTATCA 3426
QY 361 CATGTTATGAATGCTGATATAAAGACTTCATTTTAAATTTGTTGATGGCACAGTATGGA 420
DB 3425 CATGTTATGAATGCTGATATAAAGACTTCATTTTAAATTTGTTGATGGCACAGTATGGA 3366
QY 421 TTGCCTGAAAGTAGTATTTCGCTTTTGGAGCATGTGAATATAGACGTCACCTTTTAAAGT 480
DB 3365 TTGCCTGAAAGTAGTATTTCGCTTTTGGAGCATGTGAATATAGACGTCACCTTTTAAAGT 3306
QY 481 TATAAACCTGATTACGCAATTTATGACAAATATTGATTTTCGATCATCTCTGATTTATTAAA 540
DB 3305 TATAAACCTGATTACGCAATTTATGACAAATATTGATTTTCGATCATCTCTGATTTATTAAA 3246
QY 541 GATATTAATGATGTTTGTGATTCCTCAAGAAATGGCACATAATGTTTAAAAAGGTATT 600
DB 3245 GATATTAATGATGTTTGTGATTCCTCAAGAAATGGCACATAATGTTTAAAAAGGTATT 3186
QY 601 ATTGCTTTGGGTGATGATGAACATCTAGTAAAAATTGAAGCAGATGTTTCCAAATTTATTAT 660
DB 3185 ATTGCTTTGGGTGATGATGAACATCTAGTAAAAATTGAAGCAGATGTTTCCAAATTTATTAT 3126
QY 661 TATGGATTTAAAGATTGCGATGACATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 3125 TATGGATTTAAAGATTGCGATGACATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3066
QY 721 ACTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 3065 ACTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3006
QY 781 GGTGACCATACAGTTTAAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 3005 GGTGACCATACAGTTTAAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2946
QY 841 GATGTTTCAAATATTAAGAGACATTAAGAAAGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 900
DB 2945 GATGTTTCAAATATTAAGAGACATTAAGAAAGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2886
QY 901 GAAACTCAATTCGAAATCAAGTTTATGTAAGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 2885 GAAACTCAATTCGAAATCAAGTTTATGTAAGATGATGATGATGATGATGATGATGATGATGATGAT 2826
QY 961 AGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAGTTTGTGTCAGTATTTT 1020
```

```
||||| 2825 AGGCTACAAATGAAACAGACGACGAAAGAAATATCCACATAAGAAGTGTGTCAGTATTT 2766
||||| 1021 CAACACACACTTCTCTAGAACACAGCGCAATTTTAAATGAAATGTCAGAAAGTTTAAAGT 1080
||||| 2765 CAACACACACTTCTCTAGAACACAGCAATTTTAAATGAAATGTCAGAAAGTTTAAAGT 2706
||||| 1081 AAAGCAGATCGTGTATCTTATGTAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1140
||||| 2705 AAAGCAGATCGTGTATCTTATGTAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 2646
||||| 1141 TTAACGATACAGATTAATGATAAATTAAGAGTGCATCGTTAAATTAATGAAGATCT 1200
||||| 2645 TTAACGATACAGATTAATGATAAATTAAGAGTGCATCGTTAAATTAATGAAGATCT 2586
||||| 1201 ATTAATGATTTAGAACAAATTTGATGCTGTTATTTTATTTATGAGGTCAGAGTGATTT 1260
||||| 2585 ATTAATGATTTAGAACAAATTTGATGCTGTTATTTTATTTATGAGGTCAGAGTGATTT 2526
||||| 1261 CAAAAATTACAAATGCATATTTAGATAAATTTAGGCATGAAAAATGCGTTTAA 1314
||||| 2525 CAAAAATTACAAATGCATATTTAGATAAATTTAGGCATGAAAAATGCGTTTAA 2472
|||||
AF034076 1314 bp DNA linear BCT 16-JUL-1999
Staphylococcus aureus UDP-N-acetylmuramoyl-L-alanine synthetase
(murC) gene, complete cds.
AF034076.1 GI:2642658
Staphylococcus aureus
Staphylococcus aureus
Bacteria; Firmicutes; Bacillales; Staphylococcus.
Lowe, A.M. and Deresiewicz, R.L.
Cloning and sequencing of Staphylococcus aureus murC, a gene
essential for cell wall biosynthesis
DNA Seq. 10 (1), 19-23 (1999)
20029270
MEDLINE
PUBMED
10565540
Lowe, A.M. and Deresiewicz, R.L.
Direct Submission
Submitted (11-NOV-1997) Channing Laboratory, Brigham and Women's
Hospital and Harvard Medical School, 181 Longwood Ave, Boston, MA
02115, USA
FEATURES
source
1. .1314
/organism="Staphylococcus aureus"
/mol_type="genomic DNA"
/db_xref="taxon:1280"
1. .1314
/gene="murC"
1. .1314
/function="cell wall biosynthesis"
/notes="MurC; UDP-N-acetylmuramate-alanine ligase"
/codon_start=1
/transl_table=11
/product="UDP-N-acetylmuramoyl-L-alanine synthetase"
/protein_id="AAB87090.1"
/db_xref="GI:2642659"
NGKILPFPANNIKEDWVIOGNFASHEEKARAHOMKLDVVSNDPLQOIIDQYT
SVAVTGAHGTSTTGLLSHYMNGDKTSLIGDGTGMLGPSDYFAPEACEYRSHLS
KPDYALMTNIDFPDYPFDINDVDFAPQEMANNVKKGIANGDDEHLRIEADVPI
YYFGKSDDIYAQNIQITDKGTAFDVYDGFDFYVQDHTVLNALAVIAISY
LEKLDVNTKEALTFCGVKRFNETTIANQVIVDDYAHHPREISATETARKYPHK
EVAVFQPHTFRSRTOAFNEFASLSKADRVFLCEIFGSIKRENTGALTIOQLDKIEG
ASLINESINVLEQFDNAVVLFXKGAGDIQKIQNALYDLKLGKMNAP"
```

ORIGIN

```
Query Match 98.8%; Score 1302.8; DB 1; Length 1314;
Best Local Similarity 99.5%; Pred. No. 2.8e-188;
Matches 1307; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ATGACACACTATCTTTGTCGGAAATTAAGGTTCTGGCATGAGTTTCATTAGCACAAATC 60
DB 1 ATGACACACTATCTTTGTCGGAAATTAAGGTTCTGGCATGAGTTTCATTAGCACAAATC 60
QY 61 ATGCATGATTTAGGACATGAAGTTCGAAGCATCGATATTGAGAACTACGTTATTACAGAA 120
DB 61 ATGCATGATTTAGGACATGAAGTTCGAAGCATCGATATTGAGAACTACGTTATTACAGAA 120
QY 121 GTTGCTCTTTAGAAATTAAGGGGATAAAAATTAATACCAATTTGATGCTAATAACATAAAGAA 180
DB 121 GTTGCTCTTTAGAAATTAAGGGGATAAAAATTAATACCAATTTGATGCTAATAACATAAAGAA 180
QY 181 GATATGGTAGTTATACAGGTTAATGCTTCGGAGTAGGCATGAGAAATAGTACGTGCA 240
DB 181 GATATGGTAGTTATACAGGTTAATGCTTCGGAGTAGGCATGAGAAATAGTACGTGCA 240
QY 241 CATCAATTTGAAATTAGATGTTCTGTAAGTTATAATGATTTTATAGGACAGATTATTGATCAA 300
DB 241 CATCAATTTGAAATTAGATGTTCTGTAAGTTATAATGATTTTATAGGACAGATTATTGATCAA 300
QY 301 TATACCTTCAGTAGCTGTAACCTGGTGACATGTTAAACCTTCTACAAACAGGTTTATTATCA 360
DB 301 TATACCTTCAGTAGCTGTAACCTGGTGACATGTTAAACCTTCTACAAACAGGTTTATTATCA 360
QY 361 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAAATTTGGTGATGGCACAGGTATGGGA 420
DB 361 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAAATTTGGTGATGGCACAGGTATGGGA 420
QY 421 TTGCTGAAAGTGATTTTTCGCTTTTGGAGGCATGTGAATATAGACGTCACTTTTAAAGT 480
DB 421 TTGCTGAAAGTGATTTTTCGCTTTTGGAGGCATGTGAATATAGACGTCACTTTTAAAGT 480
QY 481 TATAAACCTGATTACGCAATTTATGACAAATATTAATTCGATCATCTCTGATTTATTTTAAA 540
DB 481 TATAAACCTGATTACGCAATTTATGACAAATATTAATTCGATCATCTCTGATTTATTTAAA 540
QY 541 GATATTAATGATTTTGGATGCAATTTCCAAAGAAATGGCAATTAATGTTAAAAAGGTATT 600
DB 541 GATATTAATGATTTTGGATGCAATTTCCAAAGAAATGGCAATTAATGTTAAAAAGGTATT 600
QY 601 ATTGCTTGGGTGATGATGAACATCTACGTAATAATTAAGACAGATGTTTCAATTTATTAT 660
DB 601 ATTGCTTGGGTGATGATGAACATCTACGTAATAATTAAGACAGATGTTTCAATTTATTAT 660
QY 661 TATGGAATTTAAAGATTTCGGATGACATTTTATGCTCAAAATATTCAAATTTACGGATAAAGGT 720
DB 661 TATGGAATTTAAAGATTTCGGATGACATTTTATGCTCAAAATATTCAAATTTACGGATAAAGGT 720
QY 721 ACTGCTTTTGATGTGTATGATGGATGGTGAATTTATGATCACTTCTGCTCCACAATAT 780
DB 721 ACTGCTTTTGATGTGTATGATGGATGGTGAATTTATGATCACTTCTGCTCCACAATAT 780
QY 781 GGTGACCATACAGTTTAAATGCAATGCTGTAATTTGCGATTAATTTTAGAGAGCTA 840
DB 781 GGTGACCATACAGTTTAAATGCAATGCTGTAATTTGCGATTAATTTTAGAGAGCTA 840
QY 841 GATGTTACAAATTTAAAGAGCAATTTAGAAAAGTTTGGTGGTGTAAACCTGCTTTCAAT 900
DB 841 GATGTTACAAATTTAAAGAGCAATTTAGAAAAGTTTGGTGGTGTAAACCTGCTTTCAAT 900
QY 901 GAAACTACAAATTTGCAATCAAGTTATTGTAGATGATTATGTCACACCATCCAAAGAGAAAT 960
DB 901 GAAACTACAAATTTGCAATCAAGTTATTGTAGATGATTATGTCACACCATCCAAAGAGAAAT 960
QY 961 AGTGCTTCAATTTGAAACAGCACGAGAAAGAAATATCCACATAAAGAGTGTGTGAGTATTT 1020
DB 961 AGTGCTTCAATTTGAAACAGCACGAGAAAGAAATATCCACATAAAGAGTGTGTGAGTATTT 1020
```

QY 1021 CAACACACACTTCTCTAGACACACAGGCAATTTTAAATGAAATTTGAGAAAGTTTAAAGT 1080
Db |||||
QY 1021 CAACACACACTTCTCTAGACACACAGGCAATTTTAAATGAAATTTGAGAAAGTTTAAAGT 1080
Db |||||
QY 1081 AAAGCAGATCGGTGATCTTATGCTGAAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1140
Db |||||
QY 1081 AAAGCAGATCGGTGATCTTATGCTGAAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1140
Db |||||
QY 1141 TTAAACGATCAAGATTTAAATTTGATAAAATTTGAAGTGCATCGTAAATTTATGAAGATTC 1200
Db |||||
QY 1141 TTAAACGATCAAGATTTAAATTTGATAAAATTTGAAGTGCATCGTAAATTTATGAAGATTC 1200
Db |||||
QY 1201 ATTAATGATTTAGAACAAATTTGATAATGCTGTTATTTTATTTATGAGTGCAGGTGATTT 1260
Db |||||
QY 1201 ATTAATGATTTAGAACAAATTTGATAATGCTGTTATTTTATTTAAGGTGCAGGTGATTT 1260
Db |||||
QY 1261 CAAAAATTACAAATGCATATTTAGATAAAATTTAGGATGAAATGCGTTTAA 1314
Db |||||
QY 1261 CAAAAATTACAAATGCATATTTAGATAAAATTTAGGATGAAATGCGTTTAA 1314
Db |||||

RESULT 9
AX191741 1314 bp DNA linear PAT 15-AUG-2001
LOCUS
DEFINITION Sequence 23 from Patent WO0149775.
ACCESSION AX191741
VERSION AX191741.1 GI:15209910
KEYWORDS
SOURCE Staphylococcus aureus
ORGANISM Staphylococcus aureus
REFERENCE 1
AUTHORS Iversen, P.L.
TITLE Antisense antibacterial cell division composition and method
JOURNAL Patent: WO 0149775-A 23 12-JUL-2001;
Avi Biopharma, Inc. (US)
FEATURES
source Location/Qualifiers
1 .1314
/organism="Staphylococcus aureus"
/mol_type="unassigned DNA"
/db_xref="taxon:N280"

ORIGIN
Query Match 98.8%; Score 1302.8; DB 6; Length 1314;
Best Local Similarity 99.5%; Pred. No. 2.8e-188;
Matches 1307; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ATGACACACTATCATTTCTCGGAATTTAAAGTTCTGGCATGAGTTTCATTAGCACAAATC 60
Db |||||
QY 1 ATGACACACTATCATTTCTCGGAATTTAAAGTTCTGGCATGAGTTTCATTAGCACAAATC 60
Db |||||
QY 61 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTTAGAACTACGTTATTTACAGAA 120
Db |||||
QY 61 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTTAGAACTACGTTATTTACAGAA 120
Db |||||
QY 121 GTTGCTCTTAGAAATAAGGGGATAAAATATATACCAATTTGATCTTAATACATAAAAGAA 180
Db |||||
QY 121 GTTGCTCTTAGAAATAAGGGGATAAAATATATACCAATTTGATCTTAATACATAAAAGAA 180
Db |||||
QY 181 GATATCGGTAGTTATACAGGTAAATCGGATGAGTACGATGAGGATGAGGATGAGGATGAGG 240
Db |||||
QY 181 GATATCGGTAGTTATACAGGTAAATCGGATGAGTACGATGAGGATGAGGATGAGGATGAGG 240
Db |||||
QY 241 CATCAATTTGAAATTTAGATGTTTAAATGATTTTATGATTTTATGATTTTATGATTTTATG 300
Db |||||
QY 241 CATCAATTTGAAATTTAGATGTTTAAATGATTTTATGATTTTATGATTTTATGATTTTATG 300
Db |||||
QY 301 TATATCTAGTACTGACTGCTGACATGCTGAAATCTTACACAGGTTTATTTATCA 360
Db |||||
QY 301 TATATCTAGTACTGACTGCTGACATGCTGAAATCTTACACAGGTTTATTTATCA 360
Db |||||
QY 361 CATGTTATCAATGGTGATAAAAGCTTCAATTTTAAATTTGATGCGACAGGTATGGGA 420
Db |||||

Db 361 CATGTTATGAATGGTGATATAAAAGACTTCAATTTTAAATTTGGTGATGGCACAGGTATGGGA 420
QY |||||
Db 421 TTGCTCTGAAAGTGATTTATTTGCTTTTGGGCGATGTAATATAGACGCTCACCTTTTAAAGT 480
QY |||||
Db 421 TTGCTCTGAAAGTGATTTATTTGCTTTTGGGCGATGTAATATAGACGCTCACCTTTTAAAGT 480
QY |||||
Db 481 TATAAACTGATTTAGCAATTTATGCAAAATTTGATTTTCGATCATCTCTGATTTATTTAA 540
QY |||||
Db 481 TATAAACTGATTTAGCAATTTATGCAAAATTTGATTTTCGATCATCTCTGATTTATTTAA 540
QY |||||
Db 541 GATATTAATGATGTTTGGTGGATGCAATCTACCTAAAATTTGAAGCAGATGTTTCCAAATTTAT 600
QY |||||
Db 541 GATATTAATGATGTTTGGTGGATGCAATCTACCTAAAATTTGAAGCAGATGTTTCCAAATTTAT 600
QY |||||
Db 601 ATTGCTTTGGGCGTGGATGCAATCTACCTAAAATTTGAAGCAGATGTTTCCAAATTTAT 660
QY |||||
Db 601 ATTGCTTTGGGCGTGGATGCAATCTACCTAAAATTTGAAGCAGATGTTTCCAAATTTAT 660
QY |||||
Db 661 TATGGAATTTAAAGATTCGGATGCAATTTATGCTCAAAATTTTCAAAATTTACGGATAAAGGT 720
QY |||||
Db 661 TATGGAATTTAAAGATTCGGATGCAATTTATGCTCAAAATTTTCAAAATTTACGGATAAAGGT 720
QY |||||
Db 721 ACTGCTTTTGAATGTTGATGCTGATGCTGATGCTGATGCTGATGCTGCTCTCCCAATTT 780
QY |||||
Db 721 ACTGCTTTTGAATGTTGATGCTGATGCTGATGCTGATGCTGATGCTGCTCTCCCAATTT 780
QY |||||
Db 781 GGTGACCATACAGTTTAAATGCAATTTAGCTGTAATTTGGATTTAGTTATTTAGAGAAGCTA 840
QY |||||
Db 781 GGTGACCATACAGTTTAAATGCAATTTAGCTGTAATTTGGATTTAGTTATTTAGAGAAGCTA 840
QY |||||
Db 841 GATGTTTCAAAATTTAAAGAGCAATTTAGAAACGTTTGGTGGTGTGTTTAAACGCTGTTTCAAT 900
QY |||||
Db 841 GATGTTTCAAAATTTAAAGAGCAATTTAGAAACGTTTGGTGGTGTGTTTAAACGCTGTTTCAAT 900
QY |||||
Db 901 GAAACTCAATTTGCAATTTCAATTTATTTAGTATGATTTATGCAACATTTCCCAAGGAAAT 960
QY |||||
Db 901 GAAACTCAATTTGCAATTTCAATTTATTTAGTATGATTTATGCAACATTTCCCAAGGAAAT 960
QY |||||
Db 961 AGTGCTCAATTTGAAACAGACAGAAAGAAATATCCACATAAAGAGTTGTTTCCAGTATTT 1020
QY |||||
Db 961 AGTGCTCAATTTGAAACAGACAGAAAGAAATATCCACATAAAGAGTTGTTTCCAGTATTT 1020
QY |||||
Db 1021 CAACCCACACACTTCTCTAGAACACAGGCAATTTTAAATGAAATTTTGCAGAAAGTTTAAAGT 1080
QY |||||
Db 1021 CAACCCACACACTTCTCTAGAACACAGGCAATTTTAAATGAAATTTTGCAGAAAGTTTAAAGT 1080
QY |||||
Db 1081 AAAGCAGATCGGTGATTTCTTTATGTAAGAAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1140
QY |||||
Db 1081 AAAGCAGATCGGTGATTTCTTTATGTAAGAAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1140
QY |||||
Db 1141 TTAAACGATCAAGATTTAAATTTGATAAAATTTGAAGTGCATCGTAAATTTATGAAGATTC 1200
QY |||||
Db 1141 TTAAACGATCAAGATTTAAATTTGATAAAATTTGAAGTGCATCGTAAATTTATGAAGATTC 1200
QY |||||
Db 1201 ATTAATGATTTAGAACAAATTTGATAATGCTGTTATTTTATTTATGAGTGCAGGTGATTT 1260
QY |||||
Db 1201 ATTAATGATTTAGAACAAATTTGATAATGCTGTTATTTTATTTAAGGTGCAGGTGATTT 1260
QY |||||
Db 1261 CAAAAATTACAAATGCATATTTAGATAAAATTTAGGATGAAATGCGTTTAA 1314
QY |||||
Db 1261 CAAAAATTACAAATGCATATTTAGATAAAATTTAGGATGAAATGCGTTTAA 1314
QY |||||

RESULT 10
E35613
LOCUS
DEFINITION MurC.
ACCESSION E35613
VERSION E35613.1 GI:13019102
KEYWORDS JP 1999225773-A/1.
SOURCE unidentified
ORGANISM unclassified.

E35613 1351 bp DNA linear PAT 18-JUN-2001

Acc; E3 5613

```

REFERENCE 1 (bases 1 to 1351)
AUTHORS Nicola,G.W. and Martin,K.R.B.
TITLE Murc
JOURNAL Patent: JP 1999225773-A 1 24-AUG-1999;
SMITHKLINE BEECHAM CORP,SMITHKLINE BEECHAM CORP PUBLIC LTD CO
COMMENT OS Unidentified
PN JP 1999225773-A/1
PD 24-AUG-1999
PF 03-JUL-1998 JP 1998225115
PR 03-JUL-1997 US 60/052720
PI NICOLA G WALLIS,MARTIN K R BURNHAM
PC C12N15/09,A61K31/70,A61K38/53,A61K39/085,A61K39/395,
PC A61K39/395,
PC A61K48/00,A61K48/00,C07K16/40,C12N1/21,C12P21/02,C12P21/08, PC
C12Q1/68,
PC G01N33/566//C12N15/09,C12R1:445),(C12N1/21,C12R1:445), PC
C12N15/00,A61K37/60,
PC A61K37/64,(C12N15/00,C12R1:445)
CC Strandedness: Double;
FH Key Location/Qualifiers
FT source 1..1351 /organism='Unidentified'.
FT Location/Qualifiers
source 1..1351
/organism='unidentified'
/mol type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN

Query Match 98.1%; Score 1293.2; DB 6; Length 1351;
Best Local Similarity 99.0%; Pred. No. 7.8e-187;
Matches 1301; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 ATGACACATATCATTTTGTGCGAATTAAGGTTCTGGCATGAGTTCATTTAGCAAAATC 60
DB 22 ATGACACATATCATTTTGTGCGAATTAAGGTTCTGGCATGAGTTCATTTAGCAAAATC 81
QY 61 ATGCATGATTAGACATGAGTTTCAAGGATCGGATTTAGGATTTAGGATTTACAGAA 120
DB 82 ATGCATGATTAGACATGAGTTTCAAGGATCGGATTTAGGATTTAGGATTTACAGAA 141
QY 121 GTTGCTCTTAGAATAAGGGATGATAAATATACCATTTGATGCTTAATAACATAAAGAA 180
DB 142 GTTGCTCTTAGAATAAGGGATGATAAATATACCATTTGATGCTTAATAACATAAAGAA 201
QY 181 GATATGATGATTATACAGGTAATGCAATCGCGAGTAGCCATGAAGAAATAGTCGTGCA 240
DB 202 GATATGATGATTATACAGGTAATGCAATCGCGAGTAGCCATGAAGAAATAGTCGTGCA 261
QY 241 CATCAATGGAATTAGATGTTGTAAGTTAATGATTTTATGAGGATTTATGATCAAA 300
DB 262 CATCAATGGAATTAGATGTTGTAAGTTAATGATTTTATGAGGATTTATGATCAAA 321
QY 301 TATACCTTCAGTAGCTGTAATCGTGCATCGTAAACTTCTACAAAGGTTTATTTATCA 360
DB 322 TATACCTTCAGTAGCTGTAATCGTGCATCGTAAACTTCTACAAAGGTTTATTTATCA 381
QY 361 CATGTTATGATGTTATGATAAAGAACTTCATTTTAAATGCGATGCGACAGGATGCGA 420
DB 382 CATGTTATGATGTTATGATAAAGAACTTCATTTTAAATGCGATGCGACAGGATGCGA 441
QY 421 TTGCTGAAAGTGAATTTTTCGCTTTTGGAGCATGTAATATAGACGTCATTTTAAAT 480
DB 442 TTGCTGAAAGTGAATTTTTCGCTTTTGGAGCATGTAATATAGACGTCATTTTAAAT 501
QY 481 TATAAACCCTGATTACGAATTTATGCAAAATATGATTTTTCGATCATCTCGATTTTAA 540
DB 502 TATAAACCCTGATTACGAATTTATGCAAAATATGATTTTTCGATCATCTCGATTTTAA 561
QY 541 GATATTATGATGTTTGTGATGCAATCCAAAGAAATGGACATATGTTTAAAAAGGATTT 600
DB 562 GATATTATGATGTTTGTGATGCAATCCAAAGAAATGGACATATGTTTAAAAAGGATTT 621

```

```

QY 601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTCGAACAGATGTTCCAAATTTAT 660
DB 622 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTCGAACAGATGTTCCAAATTTAT 681
QY 661 TATGGATTTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTCAGGATAAGGT 720
DB 682 TATGGATTTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTCAGGATAAGGT 741
QY 721 ACTGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 742 ACTGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 801
QY 781 GGTGACCATACAGTTTTTAAATGCAATTTAGCTGTAATTTGCGATTTAGTTATTTAGAGAACTA 840
DB 802 GGTGACCATACAGTTTTTAAATGCAATTTAGCTGTAATTTGCGATTTAGTTATTTAGAGAACTA 861
QY 841 GATGTTTACAAATTTTAAAGAGCATTTAGAAAAGTTTGGTGGTGGTGGTGGTGGTGGTGGT 900
DB 862 GATGTTTACAAATTTTAAAGAGCATTTAGAAAAGTTTGGTGGTGGTGGTGGTGGTGGTGGT 921
QY 901 GAAACTACAAATTTGCAAAATCAAGTTTATTTAGATGATTTATGCAACCATCCAAAGAGAAAT 960
DB 922 GAAACTACAAATTTGCAAAATCAAGTTTATTTAGATGATTTATGCAACCATCCAAAGAGAAAT 981
QY 961 AGTGCTTACAAATTTGAAACAGCACGAAAGAAATATCCATATAAGAAAGTTGTTGCGAGTATTT 1020
DB 982 AGTGCTTACAAATTTGACACAGCACGAAAGAAATATCCATATAAGAAAGTTGTTGCGAGTATTT 1041
QY 1021 CAAACACACACTTTCTCTAGAACACAGGCAATTTTAAATGAAATTTGAGAGAAATTTAGT 1080
DB 1042 CAAACACACACTTTCTCTAGAACACAGGCAATTTTAAATGAAATTTGAGAGAAATTTAGT 1101
QY 1081 AAAGCAGATCGTGTATTTCTTATGTAATTTTGGATCAATTTAGAGAAATTTAGTGGCGCA 1140
DB 1102 AAAGCAGATCGTGTATTTCTTATGTAATTTTGGATCAATTTAGAGAAATTTAGTGGCGCA 1161
QY 1141 TTAACGATCAAGATTTAAATTTGATAAATTTGAGAGTTCATCGTTTAAATTTAGAGATTTCT 1200
DB 1162 TTAACGATCAAGATTTAAATTTGATAAATTTGAGAGTTCATCGTTTAAATTTAGAGATTTCT 1221
QY 1201 ATTAATGTTATTAGAACAAATTTGATAAATTTGAGAGTTCATCGTTTAAATTTAGAGATTTCT 1260
DB 1222 ATTAATGTTATTAGAACAAATTTGATAAATTTGAGAGTTCATCGTTTAAATTTAGAGATTTCT 1281
QY 1261 CAAAATTTACAAATTCATATTTAGATAAATTTAGGCAATGAAATTTAGGCTTTTAA 1314
DB 1282 CAAAATTTACAAATTCATATTTAGATAAATTTAGGCAATGAAATTTAGGCTTTTAA 1335

```

RESULT 11

```

BD178030 1351 bp DNA linear PAT 16-APR-2003
Murc.
DEFINITION BD178030
ACCESSION BD178030.1 GI:30015294
VERSION JP 2002300888-A/1.
KEYWORDS unclassified
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 1351)
AUTHORS Wallis,N.G. and Burnham,M.K.R.
TITLE Patent: JP 2002300888-A 1 15-OCT-2002;
JOURNAL SMITHKLINE BEECHAM CORP,SMITHKLINE BEECHAM PLC
COMMENT OS Unidentified
PN JP 2002300888-A/1
PD 15-OCT-2002
PF 25-DEC-2001 JP 2001391079
PR 03-JUL-1997 US 60/052720
PI NICOLA G WALLIS,MARTIN K R BURNHAM
PC C12N15/09,A61K31/7088,A61K38/43,A61K38/55,A61K45/00,A61K48/00,
PC A61P1/00,

```

```
PC A61P7/02,A61P9/00,A61P11/00,A61P13/02,A61P13/12,A61P17/00, PC
A61P17/02,
PC A61P19/02,A61P19/08,A61P21/00,A61P25/00,A61P27/00,A61P29/00,
PC A61P31/04,
PC A61P39/02,A61P43/00,C07K14/195,C12N1/15,C12N1/19,C12N1/21, PC
C12N5/10,
PC C12P21/02,C12Q1/02,C12Q1/68,G01N33/15,G01N33/50,C12N15/00, PC
C12N5/00,
CC A61K37/48,A61K37/64
CC MurC
PH Key
FT source
FT Location/Qualifiers
    1. .1351
    /organism="unidentified"
    /mol_type="genomic DNA"
    /db_xref="taxon:32644"
FEATURES
    source
ORIGIN
    Query Match      98.1%; Score 1293.2; DB 6; Length 1351;
    Best Local Similarity 99.0%; Pred. No. 7.8e-187;
    Matches 1301; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1 ATGACACATATCAATTTGTCGGAATTAAGTTCTGGCATGAGTTCAATTAGCACAAATC 60
DB 22 ATGACACATATCAATTTGTCGGAATTAAGTTCTGGCATGAGTTCAATTAGCACAAATC 81
QY 61 ATGCATGATTAGGACATGAAGTTCAAGATTCGGATATTCAGAACTACGTTATTACGAA 120
DB 82 ATGCATGATTAGGACATGAAGTTCAAGATTCGGATATTCAGAACTACGTTATTACGAA 141
QY 121 GTTGCTCTTAGAAATAAGGGATGAAATAATTAACCAATTTGATGCTTAATAACATAAAGAA 180
DB 142 GTTGCTCTTAGAAATAAGGGATGAAATAATTAACCAATTTGATGCTTAATAACATAAAGAA 201
QY 181 GATATGGTATTTATACAAGGTAAATCAATTCGCGAGTAGCCATGAAGAAATPAGTACGTGCA 240
DB 202 GATATGGTATTTATACAAGGTAAATCAATTCGCGAGTAGCCATGAAGAAATPAGTACGTGCA 261
QY 241 CATCAATTCGAATTAGATGTTGAAGTTATATTAATTTTATAGGACGATTTATGATCAA 300
DB 262 CATCAATTCGAATTAGATGTTGAAGTTATATTAATTTTATAGGACGATTTATGATCAA 321
QY 301 TATACTTCAGTAGCTTAATCTGTCGACATGTAAGAACTTCTCAACAGGTTTTATTATCA 360
DB 322 TATACTTCAGTAGCTTAATCTGTCGACATGTAAGAACTTCTCAACAGGTTTTATTATCA 381
QY 361 CATGTTATGAATGGTGATAAAAGACTTCAATTTTAAATGGTGATGGCACAGGTATGGGA 420
DB 382 CATGTTATGAATGGTGATAAAAGACTTCAATTTTAAATGGTGATGGCACAGGTATGGGA 441
QY 421 TTGCGCTGAAGTGATTTATTCGCTTTTGGGACATGTAATATAGACGTCACTTTTAAAGT 480
DB 442 TTGCGCTGAAGTGATTTATTCGCTTTTGGGACATGTAATATAGACGTCACTTTTAAAGT 501
QY 481 TATAAACCTGATTACGCAATTAAGCAATATGACAAATATGATTTCGATCATCTCGATTATTTAAA 540
DB 502 TATAAACCTGATTACGCAATTAAGCAATATGACAAATATGATTTCGATCATCTCGATTATTTAAA 561
QY 541 GATATTAAATGATGTTTATGATGATTTCCAGAAATGGCACAATATGTTTAAAAAGGTATT 600
DB 562 GATATTAAATGATGTTTATGATGATTTCCAGAAATGGCACAATATGTTTAAAAAGGTATT 621
QY 601 ATTGCTTGGGGTGATGATGAACATCTACGTAATAATTTGAAGCAGATGTTTCAATTTATTAT 660
DB 622 ATTGCTTGGGGTGATGATGAACATCTACGTAATAATTTGAAGCAGATGTTTCAATTTATTAT 681
QY 661 TATGGATTAAAGATTCCGATGACATTTATGCTCAAAATTTCAATTTACGATTAAGGT 720
DB 682 TATGGATTAAAGATTCCGATGACATTTATGCTCAAAATTTCAATTTACGATTAAGGT 741
QY 721 ACTGCTTTTATGATGTTGATGATGGTGGTGGTTTATGATCACTTCCTGCTCCACAATAT 780
```

```
DB 742 ACTCCTTTGATGTTGATGTTGATGTTGATGTTTATGATCACTTCTGCTCCACAATAT 801
QY 781 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTCGATTAGTATTATTAGAGAAGCTA 840
DB 802 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTCGATTAGTATTATTAGAGAAGCTA 861
QY 841 GATGTTTCAAAATATTAAGAAAGCATTAAGAAAGTTTGGTGGTGTAAAGCTGTTTCAAT 900
DB 862 GATGTTTCAAAATATTAAGAAAGCATTAAGAAAGTTTGGTGGTGTAAAGCTGTTTCAAT 921
QY 901 GAAACTACAATTGCAATCAAGTTTATTGATGATGATTATGCAACATCCAGAGAAAT 960
DB 922 GAAACTACAATTGCAATCAAGTTTATTGATGATGATTATGCAACATCCAGAGAAAT 981
QY 961 AGTCTCAATTTGAAACAGACGAAAGAAATATCCACATAAAGAGTTTGTTCAGTATTT 1020
DB 982 AGTCTCAATTTGACACAGACGAAAGAAATATCCACATAAAGAGTTTGTTCAGTATTT 1041
QY 1021 CAACACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAAGTTTAACT 1080
DB 1042 CAACACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAAGTTTATCT 1101
QY 1081 AAACGACATCGTATTTCTTATGTAATTTTGGATCAAATTTAGAGAAAATATCTGGCGCA 1140
DB 1102 AAACGACATCGTATTTCTTATGTAATTTTGGCTCAAATTTAGAGAAAATATCTGGCGCA 1161
QY 1141 TTAACGATACAGATTTAAATGATAAATTTGAAGTGCATCGTTAAATGAATGAAGATTTCT 1200
DB 1162 TTAACGATACAGATTTAAATGATAAATTTGAAGTGCATCGTTAAATGAAGATTTCT 1221
QY 1201 ATTAATGTTATTAGAACAAATTTGATAATCTGTTATTTTATTTATGGTGACGTGATTT 1260
DB 1222 ATTAATGTTATTAGAACAAATTTGATAATCTGTTATTTTATTTATGGTGACGTGATTT 1281
QY 1261 CAAAAATTAACAAATGCATATTATGATAAATTAGGCATGAAATGCGTTTAA 1314
DB 1282 CAAAAATTAACAAATGCATATTATGATAAATTAGGCATGAAATGCGTTTAA 1335
RESULT 12
AE016748/c 300029 bp DNA linear BCT 22-APR-2003
LOCUS Staphylococcus epidermidis ATCC 12228, section 5 of 9 of the
DEFINITION complete genome.
ACCESSION AE016748 AE015929
VERSION AE016748.1 GI:27315631
KEYWORDS
SOURCE
ORGANISM Staphylococcus epidermidis ATCC 12228
REFERENCE 1 (bases 1 to 300029)
AUTHORS Zhang Y., Ren, S., Li, H., Fu, G., Lu, L., Lu, G., Jia, J., Tu, Y.,
Qin, Z., Chen, Z. and Wen, Y.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-2002) Chinese National Human Genome Center at
Shanghai, 250 Bi Bo Road, Shanghai 201203, China
FEATURES
    source
    1. .300029
    /organism="Staphylococcus epidermidis ATCC 12228"
    /mol_type="genomic DNA"
    /strain="ATCC 12228"
    /db_xref="ATCC:12228"
    /db_xref="taxon:176280"
    complement (137. .1123)
    /locus tag="SE1168"
    complement (137. .1123)
    /locus tag="SE1168"
    /codon_start=1
    /transl_table=11
    /product="thioredoxin reductase-like protein"
    /protein_id="AA004767.1"
    /db_xref="GI:27315632"
```

```
/translation="MQTIESIIIGGPGCLSAAIQKKKGITLVIEKGNVVEIYNY
PTHQTFSSSKLIGDIPFVDSKPRNQALVYREVYVGHQNLNHFEBVLVYK
INNFAITTTGGVVECKLTWATGYGQHTLEAEGCAELPKVFHYFKGAHPYFNQVV
IIGKNSADALEKAGANVLYRGEQYPKAPWILPNFBSLVNHEKIMFENA
TVTKIHTSVYKDGQGLIEDNDVFEAMIGVHPDYDLKTIIGIDHTWEYCTAPYN
RETFETVENCYIAGVIAAGNDANTIFIENGKHYGQVITQSILTKKQTPLEI"
complement (1256..2674)
/locus_tag="SE1169"
CDS
complement (1256..2674)
/locus_tag="SE1169"
/codon_start=1
/transl_table=11
/product="elastin binding protein"
/protein_id="AAO04768.1"
/db_xref="GI:27315633"
/translation="MIELIKMEGTVSNNNFKDDPEKRNQINPDEQOTELKEDDKT
NENKEADSNLSNNNQPPRNAORRKRPERATNQSKOODDKHKNKSAKTTKG
SLDRYDEAQLQQQHDKQQQNKTEKQSDNRKDKGDAALVNGTSESPEHKSSTQN
RPGPAQQQKRKSTQSKPTNKDKKAATGAGIAGAGVAGAETSRKHNNKKQDQ
SKHSHNDEKSVKNDQKQKGAAGVAGAAAGVGAHNNKHNEKNS
NQNQYNDQSKKGGFKMILLPLIAAILIIGATAIFGMAIHNHNSKSDQKIAN
QSKKSDKQCAQKSDKSDSKDKKSDKNADDDSDSNSSNPATSTNNNDNV
ANNNSYTNQONQONNQNQQATQCSHTVYQENLYRIALQYTGEGTQANVDK
IKRANGLSSNIHNGQTLVIFQ"
complement (2795..4174)
/locus_tag="SE1170"
CDS
complement (2795..4174)
/locus_tag="SE1170"
/codon_start=1
/transl_table=11
/product="probable ATP-dependent DNA helicase RecQ"
/protein_id="AAO04769.1"
/db_xref="GI:27315634"
/translation="WLNKDLKWEFGDFRFGQBEIIDSVLHGRHTLGLPTGSGKL
CYQLPTILEPXTLIIISPLISIMDDQVMQTKLKGETHVACVHSGMDEVEKRNIRLS
KSRFYLSPETLQHPNFKYLAHNFGLIVLDEAHLSEWGYDFPHVALVGKIHHF
NSATVALTAPASQLETDLSHLKLVNQKSNRRENIQLQNFNDDEGKIHWL
PFLNSGPTITTVSSQKCLBLAOMIYDSGLTYHGDLSYQERQTVQQQFNNIDP
VIVATSGAGMINKDITVTIHELSSPSNVLQIIGRAGDKGQSAISLPQPDSP
ILETLTDDITDDDTIMEFTGNLHPDSEKILSTLNTQTPFSQALDIPHQSYQRKL
GYRMGMVTNDDQCRKYLLEFFGHPSPQCCQDSRLPTIQLNRKKVRKRSFN
EKLNNLFET"
complement (4161..5123)
/locus_tag="SE1171"
CDS
complement (4161..5123)
/locus_tag="SE1171"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAO04770.1"
/db_xref="GI:27315635"
/translation="WYNILSYAMSHAFNKYNKSYNIILTKGKSHQTFDASSQQLS
LYSLPNLKYSTFQFIQKDDFKGIQVQIHQPQVYDLSLTQTFSCIQLLIQTLSHR
KESNTPIVQYQKQLYHQVIESNQVNTIDEIYLLFNKNKYNHTFLHY
LQYVESWYTRQQLSLBEGIQSELFEREMELIDILNQLDKSTYFILSQAILSP
LNTVLSVKLIKSLNKEIAQLQNVKLNTIEDHLEMYIKGLYDITLFINKKDILE
FINYYQHRGERLAFKYEHTDWTYFQIKLVIGIERGDLIAER"
5232..5480
/locus_tag="SE1172"
CDS
5232..5480
/locus_tag="SE1172"
/codon_start=1
/transl_table=11
/product="ferredoxin"
/protein_id="AAO04771.1"
/db_xref="GI:27315636"
/translation="WAKYTIYDMTCTIACGACGAAPDIYDDRGIAVILDDNQGT
AEVPEELYMEDALEGCPTDSIKTEPFGDALKFE"
complement (5572..5676)
/locus_tag="SE1173"
CDS
complement (5572..5676)
/locus_tag="SE1173"
/codon_start=1
```

```
/transl_table=11
/product="hypothetical protein"
/protein_id="AAO04772.1"
/db_xref="GI:27315637"
/translation="MSSLVPNKQPMNHKQFIWGSIHLIADYFIVYG"
complement (5885..6424)
/locus_tag="SE1174"
CDS
complement (5885..6424)
/locus_tag="SE1174"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAO04773.1"
/db_xref="GI:27315638"
/translation="MQCKRLITISMLSAVAVLTFIKFPLFPPIPYTLTDFSDVPTL
LATFLSPAGIIVALIKNLFENFGDPVGPVAVNLAVGSLSSYYVYRKRKNR
SLYGLITGVIVTIVLSILNYFLLPLYGMIPLGDLVNNVIVISGVIPFNLKIG
IILSIPLVLRRLRHIIK"
complement (7016..8782)
/locus_tag="SE1175"
CDS
complement (7016..8782)
/locus_tag="SE1175"
/codon_start=1
/transl_table=11
/product="staphylococcal respiratory response protein
SrrB"
/protein_id="AAO04774.1"
/db_xref="GI:27315639"
/translation="MRLNSVVIKLTIIIVATVILISVSLTIFQYFTQETEN
AIKDAKRISLVEKSHNKTALQNSQTLIDGPGGLIIMKMSERDSSYKDTKKQ
MFNEIKSTFKPKVEYETQITIKNGNSQSYLLGYPMAKQAGSHYSGVFI
YKDLKIEDTNNAITIIITAIITFIATSTIFAPFLSNITPKPQLKTAQKVSSEG
YSQISVATKDEIGDISRAPNNMVEIOEHIKAISSNKIRDTLLNSMVEGLINNQ
REILSNKADDDIMRHIDDESKESIEQOIEATPESQONEYLEINTYVYFISVID
RIQTNGSGIVIRVIMTNEHNDQKQOFIANNVSHLRTPISILQYTESIVDGI
EPDETSDLAIVLDSNDEQLWNYDMRMDQVLTNIDNATRYTQAGSIKISIDSDS
MSEELQWTFIDFSDNQEQWNYDMRMDQVLTNIDNATRYTQAGSIKISIDSDS
FNILTTDGTGTIAPHLKQVDFRYKVDAAARKKQGTGLGLFKIMIEHGRID
VSELCKGFSIIRLPKQIS"
complement (8766..9491)
/locus_tag="SE1176"
CDS
complement (8766..9491)
/locus_tag="SE1176"
/codon_start=1
/transl_table=11
/product="staphylococcal respiratory response protein
SrrA"
/protein_id="AAO04775.1"
/db_xref="GI:27315640"
/translation="MTNEILIVDDEDRIRLLKILYLERESFEIHEARDGKEAYELALE
NNVACILLMLPEMDGIEVASKLRSHKDTPIIMLTAKGETNREVEGSGADDYIVK
PFSPRVLRVLRKALLRTQTANTQSEPHARDIIEFSLVIDNDARVLADQDQVLT
QMIQYVWGYKFEVKSDEPTK"
complement (9625..10362)
/locus_tag="SE1177"
CDS
complement (9625..10362)
/locus_tag="SE1177"
/codon_start=1
/transl_table=11
/product="ribosomal large subunit pseudouridine synthase
B"
/protein_id="AAO04776.1"
/db_xref="GI:27315641"
/translation="MSKEQERLQRIANGSYTSRRKAEVLITEGRVKVNGQTVTELGS
KWPSDLVVDGKIQEDKQIYLLFYKPTQVITSVSDRGRVVVTFDDESRIYPV
GRLDYTSGLLLTNDGEFTNLTMPRYHTKKKVAKLKGLMREEIKOLEKTELED
GYTPAQVAKIKQDKKEKNTLIVETISEGNRQVRMFEYFGHQVVKLTRIQFPIDL
KGLNAGEGRVLTPHEVKILRHLEANGK"
complement (10355..10897)
/locus_tag="SE1178"
CDS
complement (10355..10897)
/locus_tag="SE1178"
```

		/codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AA004777.1" /db_xref="GI:27315642" /translation="MDNIALEALLVYTSDEGLEQKQIILIDININOLEDLVSKYHS HGHTIRQIGSYVLTITTKETISTYEQVDEKRSKMLSQAMETLSIIAYNQPLTRGDI EMTRGNSDGLVTLIARGLVEAKQDVHSRSHLITDITFLNVFGIENLDALPTTEED EAMDEFFSLVNQKGSNE" complement(10890..11690) /locus_tag="SE1179" complement(10890..11690) /locus_tag="SE1179" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AA004778.1" /db_xref="GI:27315643" /translation="WYSYNNFVKVALTQMPDIEVDVYVYVKLDAFNGPLDLLHLILQ KYIEDIYDIPMKALTQYMOYVHAMQLEINVASSEYLVMASELMIKSLLPOTSIE EIEEDRELVRLEIYQKEYTEILNTMKERDLYFTKHPDLDLTHLETNESWDPN QTDLIELIYAVQRVKNRILNTPKSVIEIKETFTIQATVAQTERLKHQESFNFFSL FTFHPVEQVVTHTFLALELNSKSGIVNIKQTKQDFDDIIRGVNYSIG"	
		Query Match 59.7%; Score 918.8; DB 1; Length 300029; Best Local Similarity 81.2%; Pred. No. 9.4e-131; Matches 1067; Conservative 0; Mismatches 247; Indels 0; Gaps 0;	
QY	1	ATGACACATCATATTTTGTGCGGAATTAAGGTTCTGCGATGAGTTCATTAGCACAAATC 60	
DB	248525	ATGACACATCATATTTTGTGCGGAATTAAGGTTCTGCGATGAGTTCATTAGCACAAATC 248466	
QY	61	ATGCATCATTTAGCACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACGAA 120	
DB	248465	ATGCATGACCTCGTCATGATGATACAGGTTTCAGACATAGAAATCATAGCTTTTACGAA 248406	
QY	121	GTTGCTCTTAGAAATAAGGGGATAAAAATATTACATTTGATGCTTAATPAACATAAAGAA 180	
DB	248405	GTTGCATTAAGAAATAAGGGATTAATAATTTTACCTTTGATGCAAAATAATATTACAAA 248346	
QY	181	GATATGCTAGTTATACAGGTAATGCAATTCGGAGTAGCCATGAGAAATAGTACGTGCA 240	
DB	248345	GAATGGTGTGCATCCAGGTAATGCAATTCCTGATATCATGAGAAATGTTTAGGGCA 248286	
QY	241	CATCAATGAAATTAGATGTTGTGAAGTTATATGATTTTATAGGACAGATTAATGATCAA 300	
DB	248285	CATGAATTAAGCTTGATATATAAATAATCATGACTTCTTGTTGTCATGTTATAAATCAA 248226	
QY	301	TATACTTCAGTAGCTGTAACGTGGTGCAATGGTAAAATCTTCTACAAACAGGTTTATATCA 360	
DB	248225	TATACATCTGTTGCTGTTACAGGTGACACGCGTAAACGCTCAACACTGGTTGTTGTCA 248166	
QY	361	CATGTTATGATGCTGATATAAAGACTTCAATTTTAAATTTGGTATGGCACAGGTATGGA 420	
DB	248165	CATGTGATGAATGAGATAAAGACATCATTCCTTAATCGGTGATGGTACAGGTATGGA 248106	
QY	421	TTGCTCTGAAAGTGATATTTTCGCTTTTGGAGCATGTGAATATAGACGTCACTTTTAAAGT 480	
DB	248105	CTACAGGAAGTGACTATTTTTCGCTTTTGGAGCTTGTAATACCGTAGGCATCTCTTAAGT 248046	
QY	481	TATAAACCTGATTAACGCAATATAGCAAAATATGATTTTCGATCATCTCGATATTTTAAA 540	
DB	248045	TATCATCCAGATTAATGCAATATATGACTAATAATGATTTTGTACCATCCAGATTAATTTAAA 247986	
QY	541	GATATTATGATGTTTTTATGTCATTCAGAAATGCGACATAAATGTTTAAABAAGGTATT 600	
DB	247985	AATATCGATGATGTTTATGATGCAATTCACACATATGGCAATTAATGTTTAAAGAAGGAATT 247926	
QY	601	ATTGCTCGGGTGATGATGAACATCTACGTAAAATTTGAAGCAGATGTTCCCAATTTTATTAT 660	
DB	247925	ATTGCTCGGGGATGATGAATATTTTACGTAATTTAGATGCGATATTCGGGTTTATTAT 247866	
QY	661	TATGGATTTAAAGATTCGGATGACATTTTATGCTCAAAATATTTCAAAATTCGGATTAAGGT 720	

Db	247865	TATGCTTTTAAAGAAACAGATGACATCTATGCTTAAAAATATTTCAAAATTTACTGAAAAGGTT 247806
QY	721	ACTGCTTTTGTATGTATGTGATGGTGTATTTTATGATCACTTCTGTCTCCACAATAT 780
Db	247805	ACGCAATTTGTATGATATATATAAAGCGCAATTTTATGATCAATTTCTATCCCCACAATAT 247746
QY	781	GGTGACCATACAGTTTTTAAATGTCATAGCTGTAATTCGATTTAGTATTATTAGAGAGCTA 840
Db	247745	GGAAATCAACAATTTCTTAAACGCACTTCTGCTGTAAATAGCAATTTAGTATTATTAGAAAGCATG 247686
QY	841	GATGTTTACAAATATATAAAGAAAGCATTTAGAAACGTTTGGTGGTGTAAACGTCGTTTCAAT 900
Db	247685	AATGTAGAAAATATATAAAGAAAGCACTGATATACATTTTGGTGGTGTAAACGTCGTTTAAAT 247626
QY	901	GAACCTACAATTTGCAATCAAGTTTATTTGTAGATGATTATGCAACCATCCAAAGAGAAAT 960
Db	247625	GAACAAAAGTATCAATCAAGTAAATAGTAGCACTTATGACACCATCCAAAGAGAAAT 247566
QY	961	AGTGCTACAATTTGAAACAGCAGCAAGAAATATCCACATATAAGAAAGTGTGTTGCGATATT 1020
Db	247565	AGTGCTACAATTTGACAGCAGCGCAAAAAATATCCGCAAAAAAGATGTTGTTGCCGTTTT 247506
QY	1021	CAACCACACATCTTCTCTAGAACACAGGCATTTTAAATGAATTTTGCAAAAAGTTTAACT 1080
Db	247505	CAACCACATAGCTTCTCAAGAACTCAAGCATTTTAAACGAATTTTCTGAGAGTTTAACT 247446
QY	1081	AAAGCAGATCGTGTATTTCTTATGTGAAAATTTTGGATCAATTTAGAGAAATATCTGCGCA 1140
Db	247445	AAGCAGACCAAGTATTTTATGTGAATATTTTCGTTCAATAAGAGAAATACGGGAGAT 247386
QY	1141	TTAAGCATACAAGATTTAATTTGATATAAATGAAGGTGCAATGTTAATTAATGAAGATTCT 1200
Db	247385	TTAACTATAGAAGATTTAATCAATCGTATTTGACGGCTCGACGCTTATAGATGAAAATAGT 247326
QY	1201	ATTATGATTTAGAACAAATTTGATTAATGCTGTATTTATTTATGTTGCGTCCAGGTGATATT 1260
Db	247325	ATTGATGATTTAGAAAAATTTGATTAACGCTGTAAATTTTATTTATGCGCGCTGGTGACATT 247266
QY	1261	CAAAAAATTACAAAATGTCATATTTAGATAAATTAGGCATCAAAAAATTCGTTTTTAA 1314
Db	247265	CAAAAGTTACTAAAAGCTTATTTCCAAAATTTAGCGTAAAAAATGATTTTAA 247212
RESULT 13		
E35614	E35614	
LOCUS		660 bp DNA linear PAT 18-JUN-2001
DEFINITION	Murc.	
ACCESSION	E35614	
VERSION	E35614.1 GI:13019103	
KEYWORDS	JP 1999225773-A/2.	
SOURCE	unidentified	
ORGANISM	unclassified.	
REFERENCE	1. (bases 1 to 660)	
AUTHORS	Nicola, G.W. and Martin, K.R.B.	
TITLE	Murc	
JOURNAL	Patent: JP 1999225773-A 2 24-AUG-1999;	
COMMENT	SMITHKLINE BEECHAM CORP, SMITHKLINE BEECHAM CORP PUBLIC LTD CO	
	OS Unidentified	
	FN JP 1999225773-A/2	
	PD 24-AUG-1999	
	PF 03-JUL-1998 JP 1998225115	
	PR 03-JUL-1997 US 60/052720	
	PI NICOLA G WALLIS, MARTIN K R BURNHAM	
	PC C12N15/09, A61K31/70, A61K3	

QY 1266 ATTACAAAATGCATATTTAGCATAAATTTAGCATGAAAAATGCGTTTAA 1314
 Db |||||
 601 ATTACAAAATGCATATTTAGCATAAATTTAGCATGAAAAATGCGTTTAA 649

RESULT 15
 AEO17039/c
 LOCUS 291804 bp DNA linear BCT 30-APR-2003
 DEFINITION Bacillus anthracis str. Ames section 16 of 18 of the complete genome.

ACCESSION AEO17039
 VERSION AEO17039.1
 KEYWORDS GI:30259317

SOURCE Bacillus anthracis str. Ames
 ORGANISM Bacillus anthracis str. Ames
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.

REFERENCE 1 (bases 1 to 291804)
 AUTHORS Read,T., Peterson,S., Tourasse,N., Baillie,L., Paulsen,I., Neilson,K., Tettelin,H., Fouts,D., Eisen,J., Gill,S., Holtzapple,E., Oksad,O., Helgason,E., Rilstone,J., Wu,M., Kolonay,J., Beanan,M., Dodson,R., Brinkac,L., Gwinn,M., DeBoy,R., Madupu,R., Daugherty,S., Durkin,A., Haft,D., Nelson,W., Peterson,J., Pop,M., Khouri,H., Radune,D., Benton,J., Mahamoud,Y., Jiang,L., Hance,I., Weidman,J., Berry,K., Plaut,R., Wolf,A., Watkins,K., Nierman,W., Hazen,A., Cline,R., Redmond,C., Thwaite,J., White,O., Salzberg,S., Thomason,B., Friedlander,A., Koehler,T., Hanna,P., Kolsto,A.-B. and Fraser,C.

TITLE The genome sequence of Bacillus anthracis Ames and comparison to closely related bacteria

JOURNAL Nature 423 (6935), 81-86 (2003)
 MEDLINE 22608414
 PUBMED 12721629

REFERENCE 2 (bases 1 to 291804)
 AUTHORS Read,T., Peterson,S., Tourasse,N., Baillie,L., Paulsen,I., Neilson,K., Tettelin,H., Fouts,D., Eisen,J., Gill,S., Holtzapple,E., Oksad,O., Helgason,E., Rilstone,J., Wu,M., Kolonay,J., Beanan,M., Dodson,R., Brinkac,L., Gwinn,M., DeBoy,R., Madupu,R., Daugherty,S., Durkin,A., Haft,D., Nelson,W., Peterson,J., Pop,M., Khouri,H., Radune,D., Benton,J., Mahamoud,Y., Jiang,L., Hance,I., Weidman,J., Berry,K., Plaut,R., Wolf,A., Watkins,K., Nierman,W., Hazen,A., Cline,R., Redmond,C., Thwaite,J., White,O., Salzberg,S., Thomason,B., Friedlander,A., Koehler,T., Hanna,P., Kolsto,A.-B. and Fraser,C.

TITLE Direct Submission

JOURNAL Submitted (26-MAR-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

FEATURES
 source
 1..291804
 /organism="Bacillus anthracis str. Ames"
 /mol_type="genomic DNA"
 /strain="Ames"
 /db_xref="taxon:198094"
 complement(119..475)
 /gene="rplT"
 /locus tag="BA4817"
 complement(119..475)
 /gene="rplT"
 /locus tag="BA4817"
 /note="identified by match to PFAM protein family HMW PF00453"
 /codon_start=1
 /transl_table=11
 /product="ribosomal protein L20"
 /protein_id="AAP28506.1"
 /db_xref="GI:30259318"
 /translation="MPRVKGYTQRRKKVKIAGYGSKNTLPKVAQCQWKLMA YAFDRQKQRDPKRLWITFINAARMGISRLMHGINKAGIEVNRKMLADLAVHD EKAPAEIATVAKNNIN"
 complement(513..713)
 /gene="rpmI"
 /locus tag="BA4818"
 /gene="rpmI"

QY 1201 ATTAATGTATTAGACAAATTTGATAATGCTGTATTATTTTATTTATGCGTGCAGGTGATATT 1260
Db 101911 ACGGATGTATTAAAGAAACATAAAACGGCGTTCTCTATTTTCATGGCGGAGGACATC 101852

QY 1261 CAAAAATTACAAAATGCATATTTAGATAAATT 1292
Db 101851 CAAAAATTCGAAGCAGCTTACGTAAAGAAGT 101820

Search completed: June 6, 2004, 01:40:50
Job time : 5358 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2004, 02:44:45 ; Search time 635 Seconds
(without alignments)
8817.520 Million cell updates/sec

Title: US-10-712-713-1

Perfect score: 1318

Sequence: 1 atgacacatcatctttgt.....gaaaaatgcgtttaagctt 1318

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 50

Total number of hits satisfying chosen parameters: 35

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- N_Geneseq_29Jan04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002s.*
7: Geneseqn2003as.*
8: Geneseqn2003bs.*
9: Geneseqn2003cs.*
10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1318	100.0	1318	4	AAS00189 S. aureus
2	1314	99.7	1335	9	Add15462 Staphyloc
3	1311	99.5	1332	7	AcF72770 Staphyloc
4	1263	95.8	1335	9	Add15464 Staphyloc
5	1212	92.0	1314	7	ACA19726 Prokaryot
6	957	72.6	1335	4	AAS4558 Staphyloc
7	954	72.4	1311	4	AAS51660 Staphyloc
8	842	63.9	1351	2	AAS93650 UDP-N-ace
9	740	56.1	2424	2	AAV74703 Staphyloc
10	330	25.0	660	2	AAV80065 Partial n
11	275	20.9	619	2	AAV53479 DNA encod
12	257	19.5	257	7	AAS50351 Staphyloc
13	257	19.5	257	7	ACA17633 Prokaryot
14	225	17.1	225	4	AAS49022 Staphyloc
15	225	17.1	225	4	AAS49054 Staphyloc
16	225	17.1	225	4	AAS49060 Staphyloc
17	225	17.1	225	4	AAS48994 Staphyloc
18	225	17.1	225	7	ACA16308 Prokaryot
19	225	17.1	225	7	ACA16221 Prokaryot
20	225	17.1	225	7	ACA16304 Prokaryot
21	225	17.1	225	7	ACA16261 Prokaryot
22	206	15.6	206	4	AAS48839 Staphyloc
23	206	15.6	206	7	ACA16071 Prokaryot

RESULT 1

AAS00189
ID AAS00189 standard; DNA; 1318 BP.

XX
AC AAS00189;

XX
DT 04-JUL-2001 (first entry)

XX
DE S. aureus DNA encoding UDP-N-acetylmuramate:L-alanine ligase, MurC.

XX
KW UDP-N-acetylmuramate:L-alanine ligase; MurC; immunogen; vaccine;
antibody; wound infection; cellulitis; burn infection; eyelid infection;
food poisoning; joint infection; neonatal conjunctivitis; osteomyelitis;
skin infection; scalded skin syndrome; toxic epidermal necrosis;
Ritter's disease; Lyell's disease; toxic shock syndrome; endocarditis;
ds.

XX
OS Staphylococcus aureus.

XX
FH Key Location/Qualifiers

FT CDS 1..1314

FT /*tag= a

FT /product= "MurC"

XX
WO200116292-A2.

XX
08-MAR-2001.

XX
31-AUG-2000; 2000WO-US023773.

XX
01-SEP-1999; 99US-0151933P.

XX
(HUMA-) HUMAN GENOME SCI INC.

XX
Choi GH;

XX
WPI; 2001-183259/18.

XX
P-PSDB; AAU00828.

XX
New isolated nucleic acid for use in diagnosing Staphylococcus infections
and in vaccines for eliciting immune responses to the infections.

XX
Claim 1; Page 14; 225pp; English.

XX
The sequence encodes S. aureus MurC (UDP-N-acetylmuramate:L-alanine
ligase). The polynucleotides of the invention are used to detect
Staphylococcus nucleic acids in a biological sample from an animal for
diagnosing Staphylococcus infections. The polypeptides of the invention
are used to detect anti-Staphylococcus antibodies in a biological sample
from an animal to diagnose Staphylococcus infections. The polypeptides
are also used in vaccines to elicit protective antibodies in an animal to
a member of the Staphylococcus genus and for preventing or attenuating an
infection caused by a member of the Staphylococcus genus e.g wound
infection, cellulitis, burn infection, eyelid infection, food poisoning,
joint infection, neonatal conjunctivitis, osteomyelitis, skin infection,
scalded skin syndrome (also known as toxic epidermal necrosis, Ritter's

CC disease and Lyell's disease), toxic shock syndrome and endocarditis. The
CC polynucleotides may also be used in vaccines and for preventing or
CC attenuating a *Staphylococcus* infection. Antibodies to the polypeptides
CC may be used to purify, detect and target the polypeptides in vitro and
CC in vivo diagnostic and therapeutic methods
XX
SQ Sequence 1318 BP; 464 A; 162 C; 251 G; 441 T; 0 U; 0 Other;

Query Match 100.0%; Score 1318; DB 4; Length 1318;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACACATCATCTTTTCGGAATTAAGGTTCTGGCATGAGTTCATTAGCACAATC 60
Db 1 ATGACACACATCATCTTTTCGGAATTAAGGTTCTGGCATGAGTTCATTAGCACAATC 60

QY 61 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATGGAAGTCTGATTTTACAGAA 120
Db 61 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATGGAAGTCTGATTTTACAGAA 120

QY 121 GTTGCTCTTAGAATAAGGGGATAAATAATATTACATTTGATGCTAATAACATAAAGAA 180
Db 121 GTTGCTCTTAGAATAAGGGGATAAATAATATTACATTTGATGCTAATAACATAAAGAA 180

QY 181 GATATGCTAGTTATACAGGTAATGCAATTCGCGATGAGCAATGAAGAAATAGTACGTGCA 240
Db 181 GATATGCTAGTTATACAGGTAATGCAATTCGCGATGAGCAATGAAGAAATAGTACGTGCA 240

QY 241 CATCAATTTGAATTTAGATGTTGTAAGTTATTAATGATTTTATGACAGATTTATGATCAA 300
Db 241 CATCAATTTGAATTTAGATGTTGTAAGTTATTAATGATTTTATGACAGATTTATGATCAA 300

QY 301 TATATCTTCACTAGTGTAACTGGTGACATGCTGTAATCTTCAACAGGTTTATATCA 360
Db 301 TATATCTTCACTAGTGTAACTGGTGACATGCTGTAATCTTCAACAGGTTTATATCA 360

QY 361 CATGTTATGAATCGTGATAAAGACTTCAATTTTAAATGGTGATGGCAGGATGGGA 420
Db 361 CATGTTATGAATCGTGATAAAGACTTCAATTTTAAATGGTGATGGCAGGATGGGA 420

QY 421 TTGCTCTGAAGTCAATTTTGGCTTTTGGGATGTAATAGACGTCACCTTTTAAAGT 480
Db 421 TTGCTCTGAAGTCAATTTTGGCTTTTGGGATGTAATAGACGTCACCTTTTAAAGT 480

QY 481 TATAAACCTGATTAACGCAATTAATGACAAATATGATTTGATTCGATCTCTGATTTTAA 540
Db 481 TATAAACCTGATTAACGCAATTAATGACAAATATGATTTGATTCGATCTCTGATTTTAA 540

QY 541 GATATTAAATGATTTTGTGATGCAATTCAGAAATGGCACAATAATGTTAAAGGATTT 600
Db 541 GATATTAAATGATTTTGTGATGCAATTCAGAAATGGCACAATAATGTTAAAGGATTT 600

QY 601 ATTGCTTTGGGATGATGATGACATCTACGTAATAATTTGAAGCAGATGTTCCAAATTTAT 660
Db 601 ATTGCTTTGGGATGATGATGACATCTACGTAATAATTTGAAGCAGATGTTCCAAATTTAT 660

QY 661 TATGGATTTTAAAGATTCGGATGATGATGCTCAAAATATTCGATTAAGGTAAGGTT 720
Db 661 TATGGATTTTAAAGATTCGGATGATGATGCTCAAAATATTCGATTAAGGTAAGGTT 720

QY 721 ACTGCTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 721 ACTGCTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780

QY 781 GGTGACATACAGTTTAAATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 781 GGTGACATACAGTTTAAATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 840

QY 841 GATGTTTACAAATTAAGAGCAATTAAGAACTGTTTGGTGGTGTAAACGTCGTTTCAAT 900
Db 841 GATGTTTACAAATTAAGAGCAATTAAGAACTGTTTGGTGGTGTAAACGTCGTTTCAAT 900

QY 901 GAAACTACAAATTCGAATCAAGTTTATGATGATGATGATGATGATGATGATGATGATGAT 960

Db 901 GAAACTACAAATTCGAATCAAGTTTATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 AGTGTCTCAATTTGAAACAGCAGCAAGAAATATCCACATAAAGAGTGTGTCAGTATTT 1020
Db 961 AGTGTCTCAATTTGAAACAGCAGCAAGAAATATCCACATAAAGAGTGTGTCAGTATTT 1020

QY 1021 CAACACACACATTTCTCTAGAACACAGGCAATTTTAAATGAAATTTGCAGAAAGTTTAAAGT 1080
Db 1021 CAACACACACATTTCTCTAGAACACAGGCAATTTTAAATGAAATTTGCAGAAAGTTTAAAGT 1080

QY 1081 AAAGCAGATCGTGTATTTCTTATGTAATTTTGAATTTTGAATTTGATCAATTTAGAGAAATATCTGGCGCA 1140
Db 1081 AAAGCAGATCGTGTATTTCTTATGTAATTTTGAATTTTGAATTTGATCAATTTAGAGAAATATCTGGCGCA 1140

QY 1141 TTAACGATACAGATTTTAAATTTGATATAAATTTGAAGGTGTCATGTTAATTTAATGAAGATTTCT 1200
Db 1141 TTAACGATACAGATTTTAAATTTGATATAAATTTGAAGGTGTCATGTTAATTTAATGAAGATTTCT 1200

QY 1201 ATTAATGTTATAGAACAAATTTGATTAATGCTGTTATTTTATTTATGGTGCAGGTGATATT 1260
Db 1201 ATTAATGTTATAGAACAAATTTGATTAATGCTGTTATTTTATTTATGGTGCAGGTGATATT 1260

QY 1261 CAAAAATTTACAAATGCAATTTTAGATTAATTTAGGATGAAAAATCGGTTTAAAGCTT 1318
Db 1261 CAAAAATTTACAAATGCAATTTTAGATTAATTTAGGATGAAAAATCGGTTTAAAGCTT 1318

RESULT 2
ADD15462
ID ADD15462 standard; DNA; 1335 BP.
XX
AC ADD15462;
XX
DT 15-JAN-2004 (first entry)
XX
DE Staphylococcus aureus murC DNA (SeqID 27).
XX
KW microbial; antimicrobial; membrane biosynthesis; pathogenic;
KW immunological response; vaccination; surface disinfectant;
KW personal hygiene application; food preservative; ds; gene; murC;
KW UDP-N-acetylmuramate-alanine ligase.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT CDS 1..1335
FT tag= a
FT /product= "MurC protein"
PN WO2003025007-A2.
XX
PD 27-MAR-2003.
XX
PF 20-SEP-2002; 2002WO-CA001428.
XX
PR 21-SEP-2001; 2001US-0323992P.
PR 21-SEP-2001; 2001US-0324152P.
PR 25-SEP-2001; 2001US-0324692P.
PR 26-OCT-2001; 2001US-0339924P.
PR 29-OCT-2001; 2001US-0350973P.
PR 30-OCT-2001; 2001US-0340924P.
PR 27-NOV-2001; 2001US-0333666P.
PR 18-DEC-2001; 2001US-0341732P.
PR 18-DEC-2001; 2001US-0341776P.
PR 19-DEC-2001; 2001US-0341949P.
XX
PA (AFFI-) AFFINIUM PHARM INC.
XX
ED Edwards A, Dharamsi A, Vedadi M, Alam MZ, Awrey D, Beattie B;
PI Domagala M, Houston S, Kanagarajah D, Nethery K, Ng I, Mansoury K;
PI McDonauld M, Pinder B, Viola C, Wrezel O;
XX

DR WPI; 2003-469119/44.
DR P-PSDB; ADD15463.
XX Novel crystallized recombinant polypeptides from *Staphylococcus aureus*,
PT *Streptococcus pneumoniae* and *Escherichia coli* and which are involved in
PT membrane biosynthesis, useful as targets for pathogenic bacteria.
XX Claim 20; SEQ ID NO 27; 325pp; English.
XX This invention relates to the structural and functional characterisation
CC of microbial polypeptides from *Staphylococcus aureus* (S. aureus),
CC *Streptococcus pneumoniae* (S. pneumoniae) and *Escherichia coli* (E. coli)
CC that provide novel antimicrobial targets. Specifically, it refers to
CC polypeptides that are involved in membrane biosynthesis, which play a
CC critical role in the life cycle and viability of their pathogenic species
CC of origin, and hence provide valuable drug targets. Furthermore, the
CC invention describes modified version of these proteins that facilitate
CC characterisation by labelling with isotopic or heavy atoms, and also
CC fusion proteins. These proteins provide structural and functional
CC information to aid the discovery of therapeutic molecules to treat
CC disorders associated with a particular pathogenic species. As such, they
CC are useful for inducing an immunological response in an individual and as
CC an antigen for vaccination purposes. The polypeptides are also useful for
CC developing antimicrobial agents for use as surface disinfectants,
CC personal hygiene applications and as food preservatives or in treating
CC food products to eliminate potential pathogens. This polynucleotide
CC sequence is DNA predicted from the genomic sequence of *S. aureus* UDP-N-
CC acetyl-muramate-alanine ligase (murC) of the invention.
XX
XX
SQ Sequence 1335 BP; 472 A; 161 C; 255 G; 447 T; 0 U; 0 Other;
Query Match 99.7%; Score 1314; DB 9; Length 1335;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACACATCATCTTTTGTGCGGAATTAAGGTTCTGGCATGAGTTCAATAGCAAAATC 60
DB 22 ATGACACATCATCTTTTGTGCGGAATTAAGGTTCTGGCATGAGTTCAATAGCAAAATC 81
QY 61 ATGATGATTTAGACATGAAGTTCAAGGATCGGATATGAGAACTAGTATTTACAGAA 120
DB 82 ATGATGATTTAGACATGAAGTTCAAGGATCGGATATGAGAACTAGTATTTACAGAA 141
QY 121 GTTGCTCTTGAATAAGGGATGATAAATAATATACATTTGATGCTTAATAACATAAAGAA 180
DB 142 GTTGCTCTTGAATAAGGGATGATAAATAATATACATTTGATGCTTAATAACATAAAGAA 201
QY 181 GATATGATGATTTACAGGTAATGATGCTGCGAGTAGCCATGAAGAAATAGTACGTGCA 240
DB 202 GATATGATGATTTACAGGTAATGATGCTGCGAGTAGCCATGAAGAAATAGTACGTGCA 261
QY 241 CATCAATTTGAATAGATGTTGTAAGTTATATGATTTTATAGGACAGATTTAGTATCA 300
DB 262 CATCAATTTGAATAGATGTTGTAAGTTATATGATTTTATAGGACAGATTTAGTATCA 321
QY 301 TATACCTTCAGTACGTGTAATGCTGGTGCATGCTGTAAGTTTATACACAGGTTTATATCA 360
DB 322 TATACCTTCAGTACGTGTAATGCTGGTGCATGCTGTAAGTTTATACACAGGTTTATATCA 381
QY 361 CATGTTATGATGTTGATTAAGAAAGACTTCATTTTATTTGATGGCAGGATGCGGA 420
DB 382 CATGTTATGATGTTGATTAAGAAAGACTTCATTTTATTTGATGGCAGGATGCGGA 441
QY 421 TTGCTGAAAGTGAATTTTTCGCTTTTGGGCGATGTAATATAGACGTCACTTTTAAAGT 480
DB 442 TTGCTGAAAGTGAATTTTTCGCTTTTGGGCGATGTAATATAGACGTCACTTTTAAAGT 501
QY 481 TATAACTGATTTAGCAATTTATGACAAATTTGATTTGATTCATCTCGATATTTTAAA 540
DB 502 TATAAACCCTGATTAAGCAATTTATGACAAATTTGATTTGATTCATCTCGATATTTTAAA 561
QY 541 GATATTAATGATGTTTGTGATGCTATTCAGAAATGGCAGATTAATGTTAAAAAGGTTAT 600

DB 562 GATATTAATGATGTTTGTGATGCTATTCAGAAATGGCAGATTAATGTTAAAAAGGTTAT 621
QY 601 ATTGCTTGGGTTGATGATGAACATCTACGTAAAAATGAAGCAGATGTTCCAAATTTATAT 660
DB 622 ATTGCTTGGGTTGATGATGAACATCTACGTAAAAATGAAGCAGATGTTCCAAATTTATAT 681
QY 661 TATGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTTAGGATAAGGTT 720
DB 682 TATGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTTAGGATAAGGTT 741
QY 721 ACTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 742 ACTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 801
QY 781 GGTGACCATACACAGTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 802 GGTGACCATACACAGTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 861
QY 841 GATGTTACAAATATTAAAGACATTTAGAAACGTTTGGTGGTGTAAACGTTTCAAT 900
DB 862 GATGTTACAAATATTAAAGACATTTAGAAACGTTTGGTGGTGTAAACGTTTCAAT 921
QY 901 GAACTACAAATTTGCAATCAAGTTATTAGATGATGATGATGATGATGATGATGATGATGAT 960
DB 922 GAACTACAAATTTGCAATCAAGTTATTAGATGATGATGATGATGATGATGATGATGATGAT 981
QY 961 AGTGCTTACAAATTTGAAACACACGAAAGAAATATCCACATAAAGAAAGTTGTTGCAATTT 1020
DB 982 AGTGCTTACAAATTTGAAACACACGAAAGAAATATCCACATAAAGAAAGTTGTTGCAATTT 1041
QY 1021 CAACACACACATTTCTCTAGAACACACAGGCAATTTTAAATGAAATTTGCAGAAAGTTTAAAGT 1080
DB 1042 CAACACACACATTTCTCTAGAACACACAGGCAATTTTAAATGAAATTTGCAGAAAGTTTAAAGT 1101
QY 1081 AAAGCAGATCGGTATCTTATGTAATTTTGGATCAATTTAGAGAAATATACATGCGGCA 1140
DB 1102 AAAGCAGATCGGTATCTTATGTAATTTTGGATCAATTTAGAGAAATATACATGCGGCA 1161
QY 1141 TTAAAGTACAGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1162 TTAAAGTACAGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1221
QY 1201 ATTAATGATTTAGAACAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1222 ATTAATGATTTAGAACAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1281
QY 1261 CAAAATTTACAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1314
DB 1282 CAAAATTTACAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1335
RESULT 3
ACF72770
ID ACF72770 standard; DNA; 1332 BP.
XX ACF72770;
AC ACF72770;
XX AC ACF72770;
DT 20-NOV-2003 (first entry)
XX Staphylococcus aureus DNA #450.
DE DE
KW Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW enzymatic assay; antibiotic target; gene; ds.
OS Staphylococcus aureus.
XX WO200294868-A2.
XX 28-NOV-2002.
XX 27-MAR-2002; 2002WO-IB002637.
XX 27-MAR-2001; 2001GB-00007661.
PR

XX PA (CHIR-) CHIRON SPA.
 XX PI Masignani V, Mora M, Scarselli M;
 XX DR WPI; 2003-120786/11.
 XX DR P-PSDB; ABM71210.
 XX PT New Staphylococcus aureus protein, useful as a vaccine for treating or
 XX FT preventing Staphylococcal infection, specifically an infection caused by
 XX PT S. aureus, e.g. sepsis.
 XX PS Claim 6; SEQ ID NO 899; 49pp; English.
 XX CC The invention relates to novel genes and encoded proteins from
 CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
 CC nucleic acid encoding the protein, or an antibody to the protein, is
 CC useful as a pharmaceutical, particularly as a vaccine for treating or
 CC preventing infection due to Staphylococcus bacteria, specifically an
 CC infection caused by S. aureus. The composition is particularly useful for
 CC treating or preventing sepsis in a patient. The composition can also be
 CC used for diagnostics. The protein is also used in an assay for enzymatic
 CC studies and as a target for antibiotics. This sequence represents one of
 CC the novel S. aureus genes of the invention
 XX SQ Sequence 1332 BP; 470 A; 161 C; 255 G; 446 T; 0 U; 0 Other;
 Query Match 99.5%; Score 1311; DB 7; Length 1332;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGACACACATCATTTTTCGGAATTAAGGTTCTGGCATGAGTTTCATTAGCACAAATC 60
 DB 22 ATGACACACATCATTTTTCGGAATTAAGGTTCTGGCATGAGTTTCATTAGCACAAATC 81
 QY 61 ATGCATGATTAGGACATGATGTTCAAGGATCGGATATGAGAACTAGTATTTACAGAA 120
 DB 82 ATGCATGATTAGGACATGATGTTCAAGGATCGGATATGAGAACTAGTATTTACAGAA 141
 QY 121 GTTGCTCTTAGAATAAGGGGATAAATAATTAACATTTGATGCTTAATACATAAAGAA 180
 DB 142 GTTGCTCTTAGAATAAGGGGATAAATAATTAACATTTGATGCTTAATACATAAAGAA 201
 QY 181 GATATGATGATTATACAGGTAATGCAATTCGCGAGTAGCCATGAAGAAATAGTACGTGCA 240
 DB 202 GATATGATGATTATACAGGTAATGCAATTCGCGAGTAGCCATGAAGAAATAGTACGTGCA 261
 QY 241 CATCAATTGAAATAGATGTTGTAAGTTAATGATTTTATGACAGATTAATTGATCAA 300
 DB 262 CATCAATTGAAATAGATGTTGTAAGTTAATGATTTTATGACAGATTAATTGATCAA 321
 QY 301 TATACCTTCAGTAGCTGTAACCTGGTGCACATGTTAAACTTCTACACAGGTTTATTATCA 360
 DB 322 TATACCTTCAGTAGCTGTAACCTGGTGCACATGTTAAACTTCTACACAGGTTTATTATCA 381
 QY 361 CATGTTATGAATCGGTGATATAAAGACTTCATTTTAAATGGTGGATGCAAGGTATCGGA 420
 DB 382 CATGTTATGAATCGGTGATATAAAGACTTCATTTTAAATGGTGGATGCAAGGTATCGGA 441
 QY 421 TTGCTGAAAGTGAATTAATTCGCTTTTGGGATGTAAGTATGACAGTCACTTTTAAAGT 480
 DB 442 TTGCTGAAAGTGAATTAATTCGCTTTTGGGATGTAAGTATGACAGTCACTTTTAAAGT 501
 QY 481 TATAAACCTGATTACCAATTAATGACAAATTAATGATTTTCGATCATCTCGATTATTTAAA 540
 DB 502 TATAAACCTGATTACCAATTAATGACAAATTAATGATTTTCGATCATCTCGATTATTTAAA 561
 QY 541 GATATTAAATGATGTTTTCGATGATTTCCAGAAATGACATATGTTTAAAGAGGTATT 600
 DB 562 GATATTAAATGATGTTTTCGATGATTTCCAGAAATGACATATGTTTAAAGAGGTATT 621
 QY 601 ATTGCTTGGGTGATGATGAACATCTACGTTAAATTTGAAGCAGATGTTTCAATTTATTAT 660

DB 622 ATTGCTTGGGTGATGATGAACATCTACGTAAATTTGAAGCAGATGTTTCAATTTATTAT 681
 QY 661 TATGATTTTAAAGATTTCGGATGACATTTATGCTCAAAATATTCAAATTTACGGATAAAGT 720
 DB 682 TATGATTTTAAAGATTTCGGATGACATTTATGCTCAAAATATTCAAATTTACGGATAAAGT 741
 QY 721 ACTGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 DB 742 ACTGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 801
 QY 781 GGTGACCATACAGTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 DB 802 GGTGACCATACAGTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 861
 QY 841 GATGTTTCAAAATTTAAAGCAAGCATTAGAAACGTTTGGTGGTGTAAACGTCCTTTCAAT 900
 DB 862 GATGTTTCAAAATTTAAAGCAAGCATTAGAAACGTTTGGTGGTGTAAACGTCCTTTCAAT 921
 QY 901 GAACTACAAATTTGCAAAATCAAGTTTATTTAGATGATGATGATGATGATGATGATGATGAT 960
 DB 922 GAACTACAAATTTGCAAAATCAAGTTTATTTAGATGATGATGATGATGATGATGATGATGAT 981
 QY 961 AGTGTCTCAATTTGAAACGACGACGAAAGAAATATCCACATATAAAGAGTTGTCAGTATTT 1020
 DB 982 AGTGTCTCAATTTGAAACGACGACGAAAGAAATATCCACATATAAAGAGTTGTCAGTATTT 1041
 QY 1021 CAACACACACATTTCTTAGAACACAGGCAATTTTAAATGAAATTTGCAGAAAGTTTAAAGT 1080
 DB 1042 CAACACACACATTTCTTAGAACACAGGCAATTTTAAATGAAATTTGCAGAAAGTTTAAAGT 1101
 QY 1081 AAAGCAGATCGTGTATTTCTTATGTGAAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1140
 DB 1102 AAAGCAGATCGTGTATTTCTTATGTGAAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1161
 QY 1141 TTAACGATACAGATTTAATTTGATTAATTTGAAGTGCATCGTTAATTAATGAAGATTTCT 1200
 DB 1162 TTAACGATACAGATTTAATTTGATTAATTTGAAGTGCATCGTTAATTAATGAAGATTTCT 1221
 QY 1201 ATTAATGATTTAGAACAAATTTGATTAATTTGATTAATTTGATTTATTTATTTATGGTGCAGTATTT 1260
 DB 1222 ATTAATGATTTAGAACAAATTTGATTAATTTGATTAATTTGATTTATTTATTTATGGTGCAGTATTT 1281
 QY 1261 CAAAAATTTACAAATTTGCATATTTAGATAAATTTAGGCATGAAAAATTCGGTTT 1311
 DB 1282 CAAAAATTTACAAATTTGCATATTTAGATAAATTTAGGCATGAAAAATTCGGTTT 1332
 RESULT 4
 ADD15464
 ID ADD15464 standard; DNA; 1335 BP.
 XX AC ADD15464;
 XX DT 15-JAN-2004 (first entry)
 XX DE Staphylococcus aureus murC DNA (SeqID 29).
 XX KW microbial; antimicrobial; membrane biosynthesis; pathogenic;
 KW immunological response; vaccination; surface disinfectant;
 KW personal hygiene application; food preservative; ds; gene; murC;
 XX UDP-N-acetylmuramate-alanine ligase.
 OS Staphylococcus aureus.
 XX Key Location/Qualifiers
 FH CDS 1..1335
 FT /*tag= a
 FT /product= "MurC protein"
 XX MO2003025007-A2.
 XX PD 27-MAR-2003.
 XX

PF 20-SEP-2002; 2002WO-CA001728.
XX 21-SEP-2001; 2001US-0323992P.
PF 21-SEP-2001; 2001US-0324152P.
PR 25-SEP-2001; 2001US-0324692P.
PR 25-SEP-2001; 2001US-0339924P.
PR 26-OCT-2001; 2001US-0339924P.
PR 29-OCT-2001; 2001US-0350973P.
PR 30-OCT-2001; 2001US-0340924P.
PR 27-NOV-2001; 2001US-0333666P.
PR 18-DEC-2001; 2001US-0341732P.
PR 18-DEC-2001; 2001US-0341776P.
PR 19-DEC-2001; 2001US-0341949P.
XX (AFFI-) AFFINIUM PHARM INC.
XX PA Edwards A, Vedadi M, Alam MZ, Awrey D, Beattie B;
XX PI Domagala M, Housken S, Kanagarajah D, Nethery K, Ng I, Mansoury K;
XX PI McDonald M, Pinder B, Viola C, Wrezel O;
XX WPI; 2003-468119/44.
DR P-PSDB; ADD15465.
XX Novel crystallized recombinant polypeptides from *Staphylococcus aureus*,
PT *Streptococcus pneumoniae* and *Escherichia coli* and which are involved in
PT membrane biosynthesis, useful as targets for pathogenic bacteria.
XX Claim 20; SEQ ID NO 29; 325pp; English.
CC This invention relates to the structural and functional characterisation
CC of microbial polypeptides from *Staphylococcus aureus* (*S. aureus*),
CC *Streptococcus pneumoniae* (*S. pneumoniae*) and *Escherichia coli* (*E. coli*)
CC that provide novel antimicrobial targets. Specifically, it refers to
CC polypeptides that are involved in membrane biosynthesis, which play a
CC critical role in the life cycle and viability of their pathogenic species
CC of origin, and hence provide valuable drug targets. Furthermore, the
CC invention describes modified version of these proteins that facilitate
CC fusion proteins. These proteins provide structural and functional
CC information to aid the discovery of therapeutic molecules to treat
CC disorders associated with a particular pathogenic species. As such, they
CC are useful for inducing an immunological response in an individual and as
CC an antigen for vaccination purposes. The polypeptides are also useful for
CC developing antimicrobial agents for use as surface disinfectants,
CC personal hygiene applications and as food preservatives or in treating
CC food products to eliminate potential pathogens. This polynucleotide
CC sequence is the experimentally predicted DNA of *S. aureus* UDP-N-
CC acetylmutamate-alanine ligase (murC) of the invention.
XX SQ Sequence 1335 BP; 472 A; 160 C; 255 G; 448 T; 0 U; 0 Other;
Query Match 95.8%; Score 1263; DB 9; Length 1335;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1313; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGACACACTATCTTTGTGCGAATTAAGGTTCTGGCATGAGTTTCATTAGCACAAATC 60
DB 22 ATGACACACTATCTTTGTGCGAATTAAGGTTCTGGCATGAGTTTCATTAGCACAAATC 81
QY 61 ATGCGATGTTTGGACATGAAGTTCAAGGATCGGATATTTAGAACTACGTTATTACAGAA 120
DB 82 ATGCGATGTTTGGACATGAAGTTCAAGGATCGGATATTTAGAACTACGTTATTACAGAA 141
QY 121 GTTGTCTTAGAATAAGGGGATAAATAATTTACCAATTTGATGCTTAATACATAAAGAA 180
DB 142 GTTGTCTTAGAATAAGGGGATAAATAATTTACCAATTTGATGCTTAATACATAAAGAA 201
QY 181 GATATGTTAGTTATACAAAGGTAATGCTTCGGAGTAGCATGAAGAAATAGTACGTGCA 240
DB 202 GATATGTTAGTTATACAAAGGTAATGCTTCGGAGTAGCATGAAGAAATAGTACGTGCA 261
QY 241 CATCAATTTGAAATTAGATGTTGTAAGTTATATGATTTTTTTAGGACAGATTTATGATCAA 300
DB 262 CATCAATTTGAAATTAGATGTTGTAAGTTATATGATTTTTTTAGGACAGATTTATGATCAA 321

RESULT 5
ACA19726
ID ACA19726 standard; DNA; 1314 BP.

QY 301 TATACCTCAGTAGCTGTAACCTGGTGCAATGGTAAAAACCTCTCAACAACAGGTTTATATCA 360
DB 322 TATACCTCAGTAGCTGTAACCTGGTGCAATGGTAAAAACCTCTCAACAACAGGTTTATATCA 381
QY 361 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAAATTTGGTGGTGCACAGGTTATGGGA 420
DB 382 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAAATTTGGTGGTGCACAGGTTATGGGA 441
QY 421 TTGCTCTGAAAGTGAATTTATTCGCTTTTGGGCGATGTAATATAGACGTCACCTTTTAAAGT 480
DB 442 TTGCTCTGAAAGTGAATTTATTCGCTTTTGGGCGATGTAATATAGACGTCACCTTTTAAAGT 501
QY 481 TATAAACCTGATTACGCAATTTATGACAAATTTATGATTTGATTCGATCATCTTGATTTTAAA 540
DB 502 TATAAACCTGATTACGCAATTTATGACAAATTTATGATTTGATTCGATCATCTTGATTTTAAA 561
QY 541 GATATTAAATGATGTTTGTGATGATTCCTCAAGAAATGCGACATATGTTAAAAAGGTTATT 600
DB 562 GATATTAAATGATGTTTGTGATGATTCCTCAAGAAATGCGACATATGTTAAAAAGGTTATT 621
QY 601 ATTGCTTTGGGTCGATGATGAACATCTACGTAAAAATTCGAAGCAGATGTTCCCAATTTATTAT 660
DB 622 ATTGCTTTGGGTCGATGATGAACATTTACGTAAAAATTCGAAGCAGATGTTCCCAATTTATTAT 681
QY 661 TATGGAATTTAAAGATTCGGATGACATTTATGCTCAAAATATTTCAAATTTACGGATAAAGGT 720
DB 682 TATGGAATTTAAAGATTCGGATGACATTTATGCTCAAAATATTTCAAATTTACGGATAAAGGT 741
QY 721 ACTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 742 ACTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 801
QY 781 GGTGACCATACAGCTTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 802 GGTGACCATACAGCTTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 861
QY 841 GATGTTTACAAATTTAAAGAAAGCAATTTAGAAACGTTTGGTGGTGTAAACGTCGTTTCAAT 900
DB 862 GATGTTTACAAATTTAAAGAAAGCAATTTAGAAACGTTTGGTGGTGTAAACGTCGTTTCAAT 921
QY 901 GAACTACAAATTCGAATCAAGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 922 GAACTACAAATTCGAATCAAGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 981
QY 961 AGTGCTACAAATTCGAAGCAGCAAGAAATATCCATTAAGAAAGTTGTTGCGATTTT 1020
DB 982 AGTGCTACAAATTCGAAGCAGCAAGAAATATCCATTAAGAAAGTTGTTGCGATTTT 1041
QY 1021 CAACCAACACTTTTCTCTAGAACACAGGCAATTTTAAATGAAATTTGCAAGAAAGTTTAAAGT 1080
DB 1042 CAACCAACACTTTTCTCTAGAACACAGGCAATTTTAAATGAAATTTGCAAGAAAGTTTAAAGT 1101
QY 1081 AAAGCAGATCGTGTATTCTTTATGTGAATTTTGGATCAATTTAGAGAAATATCTGGGCGCA 1140
DB 1102 AAAGCAGATCGTGTATTCTTTATGTGAATTTTGGATCAATTTAGAGAAATATCTGGGCGCA 1161
QY 1141 TTAACGATACAGATTTAAATTTGATTAATTTGAAGTGCATCGTTTAAATTAATGAAGATTTCT 1200
DB 1162 TTAACGATACAGATTTAAATTTGATTAATTTGAAGTGCATCGTTTAAATTAATGAAGATTTCT 1221
QY 1201 ATTAATGTTATAGAACAAATTTGATTAATTTGATTAATTTGATTAATTTGATTAATTTGATTAAT 1260
DB 1222 ATTAATGTTATAGAACAAATTTGATTAATTTGATTAATTTGATTAATTTGATTAATTTGATTAAT 1281
QY 1261 CAAAAATTTACAAATTCATATTTTAGATAAATTTAGGCAATGAAATATGCGTTTAA 1314
DB 1282 CAAAAATTTACAAATTCATATTTTAGATAAATTTAGGCAATGAAATATGCGTTTAA 1335

XX AC19726;
XX 19-JUN-2003 (first entry)
XX Prokaryotic essential gene #1383.
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX Staphylococcus aureus.
XX W0200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KU, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX P-PSDB; ABU15856.
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX Claim 14; SEQ ID NO 7596; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
XX prokaryotic essential genes. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 1314 BP; 464 A; 161 C; 250 G; 439 T; 0 U; 0 Other;
XX Query Match 92.0%; Score 1212; DB 7; Length 1314;

Best Local Similarity 99.8%; Pred. No. 0;
Matches 1312; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGACACATATCATTTCTCGGAATTAAGGTTCTGGCATGAGTTTCATTAGCAAAATC 60
DB 1 ATGACACATATCATTTCTCGGAATTAAGGTTCTGGCATGAGTTTCATTAGCAAAATC 60
QY 61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTCGAACTACGATTATTACAGAA 120
DB 61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTCGAACTACGATTATTACAGAA 120
QY 121 GTTGCTCTTAGAAATAAGGGGATAAAAAATATTACCAATTTGATGCTAATAACATAAAGAA 180
DB 121 GTTGCTCTTAGAAATAAGGGGATAAAAAATATTACCAATTTGATGCTAATAACATAAAGAA 180
QY 181 GATATGGTAGTTATACAAGTAAATGCAATTCGCGAGTAGCCATCAAGAAATAGTACGTCA 240
DB 181 GATATGGTAGTTATACAAGTAAATGCAATTCGCGAGTAGCCATCAAGAAATAGTACGTCA 240
QY 241 CATCAATTTGAAATTTAGATGTTGTAAGTTAATGATTTTATAGGACAGATTATTGATCAA 300
DB 241 CATCAATTTGAAATTTAGATGTTGTAAGTTAATGATTTTATAGGACAGATTATTGATCAA 300
QY 301 TATACCTCAGTAGCTGTAACTGGTGCAATGGTAAACCTTCTCAACAGGTTTATTATCA 360
DB 301 TATACCTCAGTAGCTGTAACTGGTGCAATGGTAAACCTTCTCAACAGGTTTATTATCA 360
QY 361 CATGTTATGATGGTGATATAAAGAGCTTCAATTTTAAATTTGGTATGGCAGGATATGGGA 420
DB 361 CATGTTATGATGGTGATATAAAGAGCTTCAATTTTAAATTTGGTATGGCAGGATATGGGA 420
QY 421 TTGCTCTGAAAGTGAATTTTCTGCTTTTGGAGGATGTAATATAGACGTCACTTTTAAAGT 480
DB 421 TTGCTCTGAAAGTGAATTTTCTGCTTTTGGAGGATGTAATATAGACGTCACTTTTAAAGT 480
QY 481 TATAAACCTGATTACGCAATTTATGACAAATTTGATTTTCGATCATCTGATTTTATAA 540
DB 481 TATAAACCTGATTACGCAATTTATGACAAATTTGATTTTCGATCATCTGATTTTATAA 540
QY 541 GATATTAATGATGTTTGTGATGATTCGAAATTTGCGACATAATGTTTAAAAAGGTATT 600
DB 541 GATATTAATGATGTTTGTGATGATTCGAAATTTGCGACATAATGTTTAAAAAGGTATT 600
QY 601 ATTGCTTGGGTGATGATGACATCTAGTAAATTTGAAGCAGAGTCTCCAAATTTATTAT 660
DB 601 ATTGCTTGGGTGATGATGACATCTAGTAAATTTGAAGCAGAGTCTCCAAATTTATTAT 660
QY 661 TATGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTTACGGATAAGGT 720
DB 661 TATGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTTACGGATAAGGT 720
QY 721 ACTGCTTTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 ACTGCTTTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 GGTGACCATACAGTTTAAATGCAATTTAGTCTGATTTGATGATGATGATGATGATGATGAT 840
DB 781 GGTGACCATACAGTTTAAATGCAATTTAGTCTGATTTGATGATGATGATGATGATGATGAT 840
QY 841 GATGTTCAAAATATTAAAGAGCATTAGAAAGGTTTGGTGGTGTAAACGTCGTTTCAAT 900
DB 841 GATGTTCAAAATATTAAAGAGCATTAGAAAGGTTTGGTGGTGTAAACGTCGTTTCAAT 900
QY 901 GAAACTCAATTTGAAATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 GAAACTCAATTTGAAATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 AGTGCTCAATTTGAAAGCAGCAAGAAATATCCACATAAAGAGTGTGTCAGTATT 1020
DB 961 AGTGCTCAATTTGAAAGCAGCAAGAAATATCCACATAAAGAGTGTGTCAGTATT 1020
QY 1021 CAACCCACACATTTCTCTAGAACACAGGCATTTTAAATGATTTGCGAAGTATTAAAGT 1080
DB 1021 CAACCCACACATTTCTCTAGAACACAGGCATTTTAAATGATTTGCGAAGTATTAAAGT 1080

Db 1021 CAACACACACTTCTCTAGAACACAGCATTTTAAATGAAATTTGCAGAAAGTTTAAGT 1080
QY 1081 AAAGCAGATCGGTATCTTATGTGAATTTTGGATCAATTAGAGAAATATCTGGCGCA 1140
Db 1081 AAAGCAGATCGGTATCTTATGTGAATTTTGGATCAATTAGAGAAATATCTGGCGCA 1140
QY 1141 TTAACGATACAGATTTAATTCATATAAATTTGAAGTGCATCGTTAATTAATGAAGATTCCT 1200
Db 1141 TTAACGATACAGATTTAATTCATATAAATTTGAAGTGCATCGTTAATTAATGAAGATTCCT 1200
QY 1201 ATTAATGTATTAGAACAAATTTGATATATCTGTTATTTATTTATGGGTGCGAGTGATATT 1260
Db 1201 ATTAATGTATTAGAACAAATTTGATATATCTGTTATTTATTTATGGGTGCGAGTGATATT 1260
QY 1261 CAAAATTTACAAAATGCGATATTAGATAAATTTAGGCATGAAAATGCGTTTAA 1314
Db 1261 CAAAATTTACAAAATGCGATATTAGATAAATTTAGGCATGAAAATGCGTTTAA 1314

RESULT 6

AAS54558

AAS54558 standard; DNA; 1335 BP.

AAS54558;

13-FEB-2002 (first entry)

Staphylococcus aureus DNA for cellular proliferation protein #870.

Antisense; ds; prokaryotic-cellular proliferation gene; antibiotic;

antibacterial; drug design.

Staphylococcus aureus.

WO200170955-A2.

27-SEP-2001.

21-MAR-2001; 2001WO-US009180.

21-MAR-2000; 2000US-0191078P.

23-MAY-2000; 2000US-0206848P.

26-MAY-2000; 2000US-0207272P.

23-OCT-2000; 2000US-0242578P.

27-NOV-2000; 2000US-0253625P.

22-DEC-2000; 2000US-0257931P.

16-FEB-2001; 2001US-0269308P.

(ELIT-) ELITRA PHARM INC.

Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

Yamamoto RT, Xu HH;

WPI; 2001-611495/70.

P-PSDB; AAD36699.

New polynucleotides for the identification and development of

antibiotics, comprise sequences of antisense nucleic acids.

Claim 27; SEQ ID NO 8195; 511pp; English.

The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous

CC nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence encodes an essential prokaryotic cellular proliferation protein. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 1335 BP; 472 A; 164 C; 255 G; 444 T; 0 U; 0 Other;

Query Match 72.6%; Score 957; DB 4; Length 1335;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1307; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGACACACTATCATTTTGTGCGAATTTAAAGGTTCTGGCATGAGTTCAATAGCACAAATC 60

Db 22 ATGACACACTATCATTTTGTGCGAATTTAAAGGTTCTGGCATGAGTTCAATAGCACAAATC 81

QY 61 ATGCATGATTAGGACATGAAAGTTCAAGGATCGGATATTGAGAACTACGTTATTACAGAA 120

Db 82 ATGCATGATTAGGACATGAAAGTTCAAGGATCGGATATTGAGAACTACGTTATTACAGAA 141

QY 121 GTTGCTCTTAAATAAGGGGATAAATAATATACCATTTGATGCTTAATAACATAAAGAA 180

Db 142 GTTGCTCTTAAATAAGGGGATAAATAATATACCATTTGATGCTTAATAACATAAAGAA 201

QY 181 GATATGGTAGTTATACAAAGGTAATGCATTCGCGAGTAGCCATGAAGAAATAGTAGTGCA 240

Db 202 GATATGGTAGTTATACAAAGGTAATGCATTCGCGAGTAGCCATGAAGAAATAGTAGTGCA 261

QY 241 CATCAATTTGAAATTAGAGTTTAAAGTTATAATGATTTTTTTAGGACAGATTATTGATCAA 300

Db 262 CATCAATTTGAAATTAGAGTTTAAAGTTATAATGATTTTTTTAGGACAGATTATTGATCAA 321

QY 301 TATACCTCAGTAGCTGTAACCTGGTGACATGGTAAAACTTCTACACAGGTTTATTATCA 360

Db 322 TATACCTCAGTAGCTGTAACCTGGTGACATGGTAAAACTTCTACACAGGTTTATTATCA 381

QY 361 CATGTTATGAATGTTGATAAAAAAGACTTCAATTTTAAATTTGGTGATGGCACAGGTTATGGA 420

Db 382 CATGTTATGAATGTTGATAAAAAAGACTTCAATTTTAAATTTGGTGATGGCACAGGTTATGGA 441

QY 421 TTGCCTGAAAGTGATTTATTTCCGCTTTTGGCGATGTGAATATAGACGTCACTTTTAAAGT 480

Db 442 TTGCCTGAAAGTGATTTATTTCCGCTTTTGGCGATGTGAATATAGACGTCACTTTTAAAGT 501

QY 481 TATAAACCTTGATACCGCAATTTATGACAAATATTGATTTGATTCGATCATCTGATTATTAAA 540

Db 502 TATAAACCTTGATACCGCAATTTATGACAAATATTGATTTGATTCGATCATCTGATTATTAAA 561

QY 541 GATATTAATGATGTTTGTGATGCAATTCGAAGAAATGGCACATAATGTTAAAAAGGTTATT 600

Db 562 GATATTAATGATGTTTGTGATGCAATTCGAAGAAATGGCACATAATGTTAAAAAGGTTATT 621

QY 601 ATTGCTTGGGGTGATGATGAACATCTACGTAATAATTTGAAGCAGATGTTTCCAAATTTATTAT 660

Db 622 ATTGCTTGGGGTGATGATGAACATCTACGTAATAATTTGAAGCAGATGTTTCCAAATTTATTAT 681

QY 661 TATGGAATTTAAAGATTCCGGATGACATTTATGCTCAAAATATTCAAATTTACGATAAAGGT 720

Db 682 TATGGAATTTAAAGATTCCGGATGACATTTATGCTCAAAATATTCAAATTTACGATAAAGGT 741

QY 721 ACTGCTTTTGTGATGTTATGTTGATGTTGATGTTTATGATCCTTCCTGCTCCACATAT 780

Db 742 ACTGCTTTTGTGATGTTATGTTGATGTTGATGTTTATGATCCTTCCTGCTCCACATAT 801

QY 781 GGTGACCATACAGTTTAAATGATTTAGCTGTAATTTGCGATTAGTTATTATTAGAGAGCTA 840

Db 802 GGTGACCATACAGTTTAAATGATTTAGCTGTAATTTGCGATTAGTTATTATTAGAGAGCTA 861

QY 841 GATGTTACAAATTTAAAGAACGATTAGAAAACGTTTGGTGGTGTAAACGTCGTTTCAAT 900

Db 862 GATGTTACAAATTTAAAGAACGATTAGAAAACGTTTGGTGGTGTAAACGTCGTTTCAAT 921

QY 901 GAACTACAAATTCGAATCAAGTTATTTAGATGATTATATCCACACCAATCCCAAGAGAATT 960
DB 922 GAACTACAAATTCGAATCAAGTTATTCGTAGATGATTATGACACCAATCCCAAGAGAATT 981
QY 961 AGTGTACAAATTTGAACAGCAGCAAGAAAGAAATATCCACATAAAGAGTTGTTGCAGTATTT 1020
DB 982 AGTGTACAAATTTGAACAGCAGCAAGAAAGAAATATCCACATAAAGAGTTGTTGCAGTATTT 1041
QY 1021 CAACACACACATTTCTCTAGAACACACAGGCAATTTTAAATGCAATTTGCAGAAAGTTTAAAT 1080
DB 1042 CAACACACACATTTCTCTAGAACACACAGCAATTTTAAATGCAATTTGCAGAAAGTTTAAAT 1101
QY 1081 AAAGCAGATCGTGATTTCTATCTGTAATTTTGAATTTTGGATCAATTAGAGAAAATCTGCGGCA 1140
DB 1102 AAAGCAGATCGTGATTTCTATCTGTAATTTTGGATCAATTAGAGAAAATCTGCGGCA 1161
QY 1141 TTAACGATACAGATTTAAATTTGATAAATTTGAAGTTGATCGTCTTAATTAATGAATCTT 1200
DB 1162 TTAACGATACAGATTTAAATTTGATAAATTTGAAGTTGATCGTCTTAATTAATGAATCTT 1221
QY 1201 ATTAATGATTTAGAACAAATTTGATAATCTGTTATTTTATTTATGCTGCGAGTGATTT 1260
DB 1222 ATTAATGATTTAGAACAAATTTGATAATCTGTTATTTTATTTATGCTGCGAGTGATTT 1281
QY 1261 CAAAATTTACAAATTCATATTTAGATAAATTTAGGATGAAAATGCGTTTAA 1314
DB 1282 CAAAATTTACAAATTCATATTTAGATAAATTTAGGATGAAAATGCGTTTAA 1335

RESULT 7
ID AAS51660
AC AAS51660 standard; DNA; 1311 BP.
XX AAS51660;
DT 13-FEB-2002 (first entry)
DE Staphylococcus aureus DNA for cellular proliferation protein #77.
KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
KW antibacterial; drug design.
OS Staphylococcus aureus.
PN WO200170955-A2.
PD 27-SEP-2001.
PF 21-MAR-2001; 2001WO-US009180.
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.

(ELIT-) ELITRA PHARM INC.
XX PA
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
PI
XX WPI; 2001-611495/70.
DR P-PSDB; AAU33801.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Claim 27; SEQ ID NO 4242; 51lpp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC

CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1311 BP; 462 A; 164 C; 250 G; 435 T; 0 U; 0 Other;
Query Match 72.4%; Score 954; DB 4; Length 1311;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ATGCACACTATCATTTTGTTCGGAATTAAGGTTCTCGCATGAGTTTCATTAGCACAAATC 60
DB 1 ATGCACACTATCATTTTGTTCGGAATTAAGGTTCTCGCATGAGTTTCATTAGCACAAATC 60
QY 61 ATGCATGATTTAGGACATGAAGTTCAAGATCGGATATTGGAATCTAGCTATTTCAGAA 120
DB 61 ATGCATGATTTAGGACATGAAGTTCAAGATCGGATATTGGAATCTAGCTATTTCAGAA 120
QY 121 GTTGCTCTTGAATAAGGGGATAAAAATATTACCATTTTGATGCTAAATAACATAAAGAA 180
DB 121 GTTGCTCTTGAATAAGGGGATAAAAATATTACCATTTTGATGCTAAATAAAGAA 180
QY 181 GATATGTTAGTTATACAAAGGTAATGCAATTCGCGAGTAGCCATGAAGAATAGTACGTGA 240
DB 181 GATATGTTAGTTATACAAAGGTAATGCAATTCGCGAGTAGCCATGAAGAATAGTACGTGA 240
QY 241 CATCAATTTGAATTTAGATGTTTGAAGTTTATTAATGATTTTGGACAGATTTATGATCA 300
DB 241 CATCAATTTGAATTTAGATGTTTGAAGTTTATTAATGATTTTGGACAGATTTATGATCA 300
QY 301 TATACCTTCAGTAGCTGTAACCTGGTGACATCGTAAACCTTCTACACAGGTTTATATCA 360
DB 301 TATACCTTCAGTAGCTGTAACCTGGTGACATCGTAAACCTTCTACACAGGTTTATATCA 360
QY 361 CATGTTATGATGTTGATGATAAAGACTTCAATTTTAAATTTGGTGATGGCAGGTATGGA 420
DB 361 CATGTTATGATGTTGATGATAAAGACTTCAATTTTAAATTTGGTGATGGCAGGTATGGA 420
QY 421 TTGCTGAAAGTGAATTTTCGCTTTTGGAGCATGTGAATATAGACGTCACCTTTTAACT 480
DB 421 TTGCTGAAAGTGAATTTTCGCTTTTGGAGCATGTGAATATAGACGTCACCTTTTAACT 480
QY 481 TATAAACCCTGATTTAGCAGATTTATGCAAAATTTGATTTTCGATCATCTCTGATTTTAA 540
DB 481 TATAAACCCTGATTTAGCAGATTTATGCAAAATTTGATTTTCGATCATCTCTGATTTTAA 540
QY 541 GATATTAATGATGTTTTCGATGCAATTCAGAAATTTGGCAGCATTAATGTTTAAAGGATTT 600
DB 541 GATATTAATGATGTTTTCGATGCAATTCAGAAATTTGGCAGCATTAATGTTTAAAGGATTT 600
QY 601 ATTGCTTTGGGTGATGATGAACATCTACCTAAAATTTGAAGCAGATGTTTCCAAATTTAT 660
DB 601 ATTGCTTTGGGTGATGATGAACATCTACCTAAAATTTGAAGCAGATGTTTCCAAATTTAT 660
QY 661 TATGATTTAAAGATTTGGATGACATTTATCTCAAATTTTCAATTTACGATAAAGGT 720
DB 661 TATGATTTAAAGATTTGGATGACATTTATCTCAAATTTTCAATTTTACGATAAAGGT 720
QY 721 ACTGCTTTTGGATGTTATGTTGGATGTTTATGATCACTTCTCTCTCCCAATAT 780
DB 721 ACTGCTTTTGGATGTTATGTTGGATGTTTATGATCACTTCTCTCTCCCAATAT 780

Db 562 GATATTATGATGTTTGGTGGTATCCAGAAATGGGCATTAATGTTTAAAAAGGTATT 621
Qy 601 ATTGCTGGGTGATGATGAACATCTACGTAAAAATTGAAGCAGATGTTTCCAAATTTATAT 660
Db 622 ATTGCTGGGTGATGATGAACATCTACGTAAAAATTGAAGCAGATGTTTCCAAATTTATAC 681
Qy 661 TATGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTCAGGATAAAGT 720
Db 682 TATGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTCAGGATAAAGT 741
Qy 721 ACTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 742 ACTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 801
Qy 781 GGTGACCATACAGTTTAAATGCAATTAGCTGTAATTCGATTTAGTATTTAGAGAGCTA 840
Db 802 GGTGACCATACAGTTTAAATGCAATTAGCTGTAATTCGATTTAGTATTTAGAGAGCTA 861
Qy 841 GATGTTACAAATATTAAGAGCAATAGAAAGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGT 900
Db 862 GATGTTACAAATATTAAGAGCAATAGAAAGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGT 921
Qy 901 GAAACTCAATTCGAAATCAAGTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 922 GAAACTCAATTCGAAATCAAGTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 981
Qy 961 AGTGCTCAATTTGAAACAGACGACGAAAGAAATATCCATAAAGAGTTGTTGGCATTTT 1020
Db 982 AGTGCTCAATTTGAAACAGACGACGAAAGAAATATCCATAAAGAGTTGTTGGCATTTT 1041
Qy 1021 CAACCCACACATTTCTCTAGAACACA 1046
Db 1042 CAACCCACACATTTCTCTAGAACACA 1067

RESULT 9
AAV74703
ID AAV74703 standard; DNA; 2424 BP.
AC AAV74703;
DT 16-MAR-1999 (first entry)
DE Staphylococcus aureus contig SEQ ID #392.
KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.
OS Staphylococcus aureus.

FH Key Location/Qualifiers
FT misc_feature 1141..1200
FT /tag= a
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They are
FT included to maintain the nucleotide numbering given in
FT the specification for this DNA sequence"
PN EP786519-A2.
XX 30-JUL-1997.
XX 07-JAN-1997; 97EP-00100117.
XX 05-JAN-1996; 96US-0009861P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;

DR WPI; 1997-374922/35.
XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT stored on computer readable medium and used in the production of anti-
PT S.aureus vaccines.
PS Claim 1; Page 1287-1288; 3271pp; English.
XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S.aureus DNA sequences contained on the computer
CC readable medium
SQ Sequence 2424 BP; 839 A; 302 C; 462 G; 759 T; 0 U; 62 Other;

Query Match 56.1%; Score 740; DB 2; Length 2424;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 575 TGGCACATAATGTTAAAAAGGTATTATTGCTGGGGTGGATGATGAACATCTACGTAA 634
Db 1201 TGGCACATAATGTTAAAAAGGTATTATTGCTGGGGTGGATGATGAACATCTACGTAA 1260
Qy 635 TTGAAGCAGATGTTCCAAATTTATTTATGATGATTTAAAGATTCGGATGATGATGCTC 694
Db 1261 TTGAAGCAGATGTTCCAAATTTATTTATGATGATTTAAAGATTCGGATGATGATGCTC 1320
Qy 695 AAAATATTCAAATTCAGGATAAGGTACTGCTTTTGTGATGATGATGATGATGATGATGAT 754
Db 1321 AAAATATTCAAATTCAGGATAAGGTACTGCTTTTGTGATGATGATGATGATGATGATGAT 1380
Qy 755 ATGATCATCTCTGCTCCCAATATGATGATGATGATGATGATGATGATGATGATGATGAT 814
Db 1381 ATGATCATCTCTGCTCCCAATATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
Qy 815 TTGCGATTTAGTATTATTAGAGAGCTAGATGTTTAAATTTAAAGAGCATTAGAAGCAT 874
Db 1441 TTGCGATTTAGTATTATTAGAGAGCTAGATGTTTAAATTTAAAGAGCATTAGAAGCAT 1500
Qy 875 TTGCTGGTGTAAACGTCGTTTCAATGAACTTACAAATTCGAATCAAGTTATTGTTAGATG 934
Db 1501 TTGCTGGTGTAAACGTCGTTTCAATGAACTTACAAATTCGAATCAAGTTATTGTTAGATG 1560
Qy 935 ATTATGACACCATCCCAAGAGAAATAGTGTCTACAAATTTGAAACAGCAGAAAGAAATATC 994
Db 1561 ATTATGACACCATCCCAAGAGAAATAGTGTCTACAAATTTGAAACAGCAGAAAGAAATATC 1620
Qy 995 CACATAAAGAGTTGTTGCGATTTTCAACACACATCTTCTCTAGAAACAGGCAATTTT 1054
Db 1621 CACATAAAGAGTTGTTGCGATTTTCAACACACATCTTCTCTAGAAACAGGCAATTTT 1680
Qy 1055 TAAATGAATTTGCAGAAAGTTTAAAGTAAGCAGATCGTGTATCTTATGTGAAATTTTG 1114
Db 1681 TAAATGAATTTGCAGAAAGTTTAAAGTAAGCAGATCGTGTATCTTATGTGAAATTTTG 1740
Qy 1115 GATCAATTAGAGAAATATCTGGCGCATTAACGATACAGATTTAATTCATTAATTTGAAG 1174
Db 1741 GATCAATTAGAGAAATATCTGGCGCATTAACGATACAGATTTAATTCATTAATTTGAAG 1800

QY 1175 GTGATCGTTAAATGAAGATCTATTATGATTAGAACAAATTTGATTAATGCTGTTA 1234
Db 1801 GTGATCGTTAAATGAAGATCTATTATGATTAGAACAAATTTGATTAATGCTGTTA 1860
QY 1235 TTTTATTATGCGTGCAGGTGATATTCAAAAATTTACAAAATGCAATATTAGATAAATTAG 1294
Db 1861 TTTTATTATGCGTGCAGGTGATATTCAAAAATTTACAAAATGCAATATTAGATAAATTAG 1920
QY 1295 GCATGAATAATGCGTTTAA 1314
Db 1921 GCATGAATAATGCGTTTAA 1940

RESULT 10
AAV80065
ID AAV80065 standard; DNA; 660 BP.
XX
AC AAV80065;
XX
DT 17-MAR-1999 (first entry)
XX
DE Partial nucleotide sequence of the MurC gene.
XX
KW MurC gene; UDP-N-acetylmuramate:L-alanine ligase; MurC polypeptide;
KW bacterial; infection; H. pylori; cancer; ulcer; gastritis; vaccine;
KW immunogen; drug; genetic immunisation; ds.
XX
OS Staphylococcus aureus.
XX
Key Location/Qualifiers
FH CDS
FT 2..660
FT /*tag= a
FT /product= "partial MurC polypeptide"

XX EP889123-A2.
XX 07-JAN-1999.
XX 26-JUN-1998; 98EP-00305064.
XX 03-JUL-1997; 97US-0052720P.
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX (SMIK) SMITHKLINE BEECHAM PLC.
XX Wallis NG, Burnham MKR;
XX WPI: 1999-062655/06.
XX P-PSDB; AAW89199.
XX
XX New isolated MurC polypeptide from Staphylococcus aureus and related
XX nucleic acid - useful in diagnosis, treatment and prevention of bacterial
XX infections.
XX
XX Claim 2; Page 4-5; 39pp; English.

XX The invention relates to a UDP-N-acetylmuramate:L-alanine ligase (MurC
XX polypeptide) encoded by the S. aureus MurC gene. Host cells containing an
XX expression system comprising the MurC gene can be used for the
XX recombinant production of the polypeptide. Agonists or the MurC
XX polypeptide are used to treat conditions requiring increased activity or
XX expression of the polypeptide. Antagonists, inhibitory nucleic acid or
XX competitive polypeptide are useful for inhibiting the polypeptide e.g.
XX bacterial (especially S. aureus) infections. They are also useful against
XX Helicobacter pylori infections and related cancers, ulcers and gastritis.
XX The antibacterial agents are useful to treat in-dwelling devices for
XX infection prevention or generally as wound treatments to prevent adhesion
XX of bacteria to matrix proteins. The MurC polypeptide is also useful for
XX diagnosing or prognosing a (susceptibility to) disease, for raising
XX antibodies; to identify modulators or specific receptors; in rational
XX drug design and as an immunogen for vaccines. The MurC gene sequences are
XX useful in antisense/ribozyme therapeutics; to detect mutant MurC gene;
XX for chromosomal mapping; to determine bacterial serotype; and for genetic

CC immunisation. The present sequence represents a partial nucleotide
CC sequence (MurC ORF) of the MurC gene
XX
SQ Sequence 660 BP; 233 A; 83 C; 120 G; 224 T; 0 U; 0 Other;
Query Match 25.0%; Score 330; DB 2; Length 660;
Best Local Similarity 99.7%; Pred. No. 2.3e-142;
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 666 ATTTAAAGATTCGGATGACATTTATGCTCAAAAATTTCAAAATACGGATAAAGGTACTGC 725
Db 1 ATTTAAAGATTCGGATGACATTTATGCTCAAAAATTTCAAAATACGGATAAAGGTACTGC 60
QY 726 TTTTGAATGATGATGCGATGGTGGTATTTATGATCACTTCCTGCTCCCAATATCGTGA 785
Db 61 TTTTGAATGATGATGCGATGGTGGTATTTATGATCACTTCCTGCTCCCAATATCGTGA 120
QY 786 CCATACAGTTTAAATGCAATGATGCTGTAATTCGGATTAGTTATTATTAGAGAAGCTAGATGT 845
Db 121 CCATACAGTTTAAATGCAATGATGCTGTAATTCGGATTAGTTATTATTAGAGAAGCTAGATGT 180
QY 846 TACAATATTAAAGAACGATTTAGAAACGTTTGGTGGTGTAAACGTCGTTTCAATGAAC 905
Db 181 TACAATATTAAAGAACGATTTAGAAACGTTTGGTGGTGTAAACGTCGTTTCAATGAAC 240
QY 906 TACAATTTGCAATCAAGTTTATTGTAGATGATTATGACACCATCCCAAGAGAAATTTAGTGC 965
Db 241 TACAATTTGCAATCAAGTTTATTGTAGATGATTATGACACCATCCCAAGAGAAATTTAGTGC 300
QY 966 TACAATTTGCAATCAAGTTTATTGTAGATGATTATGACACCATCCCAAGAGAAATTTAGTGC 1025
Db 301 TACAATTTGCAATCAAGTTTATTGTAGATGATTATGACACCATCCCAAGAGAAATTTAGTGC 360
QY 1026 ACACACTTTCTCTAGAACACA 1046
Db 361 ACACACTTTCTCTAGAACACA 381

RESULT 11
AAV53479/c
ID AAV53479 standard; DNA; 619 BP.
XX
XX AAV53479;
XX AC
XX 30-OCT-1998 (first entry)
XX
XX DNA encoding a Staphylococcus aureus protein of unknown function.
XX
XX Staphylococcus aureus protein; immune response induction; eye infection;
XX antibody production; T-cell immune response; gastrointestinal infection;
XX respiratory infection; inhibitor; bacterial infection; cardiac infection;
XX central nervous system; kidney infection; urinary tract infection;
XX antimicrobial compound identification; broad spectrum antibiotic;
XX therapy; ss.
XX
XX Staphylococcus aureus.
XX
XX EP841394-A2.
XX
XX 13-MAY-1998.
XX
XX 24-SEP-1997; 97EP-00307485.
XX
XX 24-SEP-1996; 96US-0027032P.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX (SMIK) SMITHKLINE BEECHAM PLC.
XX
XX Black MT, Hodgson JE, Knowles DJC, Reichard RW, Nicholas RO;
XX Burnham MKR, Pratt JM, Rosenberg M, Ward JM, Lonetto MA;
XX WPI: 1998-252940/23.
XX P-PSDB; AAW77686.

XX DE Prokaryotic essential gene antisense oligonucleotide #5503.
 XX KW Antisense; ss; prokaryotic essential gene; cell proliferation;
 XX NW drug design.
 XX OS Archaea.
 XX PN WO200277183-A2.
 XX PD 03-OCT-2002.
 XX PF 21-MAR-2002; 2002WO-US009107.
 XX PF 21-MAR-2001; 2001US-00815242.
 XX PR 06-SEP-2001; 2001US-00948993.
 XX PR 25-OCT-2001; 2001US-0342923P.
 XX PR 08-FEB-2002; 2002US-00072851.
 XX PR 06-MAR-2002; 2002US-0362699P.
 XX PA (ELIT-) ELITRA PHARM INC.
 XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX DR WPI; 2003-029926/02.
 XX PT New antisense nucleic acids, useful for identifying proteins or screening
 XX PT for homologous nucleic acids required for cellular proliferation to
 XX PT isolate candidate molecules for rational drug discovery programs.
 XX PS Claim 1; SEQ ID NO 5503; 1766pp; English.
 XX CC The invention relates to an isolated nucleic acid comprising any one of
 XX CC the 6213 antisense sequences given in the specification where expression
 XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
 XX CC encoding a polypeptide whose expression is inhibited by the antisense
 XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 XX CC polypeptide or its fragment whose expression is inhibited by the
 XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
 XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 XX CC proliferation or the activity of a gene in an operon required for
 XX CC proliferation; (7) identifying a compound that influences the activity of
 XX CC the gene product or that has an activity against a biological pathway
 XX CC required for proliferation, or that inhibits cellular proliferation. (8)
 XX CC identifying a gene required for cellular proliferation or the biological
 XX CC pathway in which a proliferation-required gene or its gene product lies
 XX CC or a gene on which the test compound that inhibits proliferation of an
 XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 XX CC compound's activity; (11) a culture comprising strains in which the gene
 XX CC product is overexpressed or underexpressed; (12) determining the extent
 XX CC to which each of the strains is present in a culture or collection of
 XX CC strains; or (13) identifying the target of a compound that inhibits the
 XX CC proliferation of an organism. The antisense nucleic acids are useful for
 XX CC identifying proteins or screening for homologous nucleic acids required
 XX CC for cellular proliferation to isolate candidate molecules for rational
 XX CC drug discovery programs, or for screening homologous nucleic acids
 XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the 6213
 XX CC antisense sequences of the invention. Note: The sequence data for this
 XX CC patent did not form part of the printed specification, but was obtained
 XX CC in electronic format directly from WIPO at
 XX CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 27-OCT-2003 to
 XX CC standardise OS field)
 XX SQ Sequence 257 BP; 84 A; 50 C; 31 G; 92 T; 0 U; 0 Other;
 XX Query Match 19.5%; Score 257; DB 7; Length 257;
 XX Best Local Similarity 100.0%; Pred. No. 1.4e-108;
 XX Matches 257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX QY 153 ACCATTGGTCTAATAACATAAAGAGATATGCTAGTATACAGGTAAATGCATTCGC 212

Db 257 ACCATTGGTCTAATAACATAAAGAGATATGCTAGTATACAGGTAAATGCATTCGC 198
 QY 213 GAGTAGCCATGACGAAGAATAGTACGTGCACATCAATGGAATAGATGTTGTAAGTTATAA 272
 Db 197 GAGTAGCCATGACGAAGAATAGTACGTGCACATCAATGGAATAGATGTTGTAAGTTATAA 138
 QY 273 TGATTTTATAGGACAGATATTGATCAATATATCACTTCACTAGCTGTAACCTGGTGACATGG 332
 Db 137 TGATTTTATAGGACAGATATTGATCAATATATCACTTCACTAGCTGTAACCTGGTGACATGG 78
 QY 333 TAAAACTTCTACACAGGTTTATTATCACTGTTATGAATGGTGATAAAAAGACTTCATT 392
 Db 77 TAAAACTTCTACACAGGTTTATTATCACTGTTATGAATGGTGATAAAAAGACTTCATT 18
 QY 393 TTTAATTGGTGATGGCA 409
 Db 17 TTTAATTGGTGATGGCA 1

RESULT 14
 AAS49022/c
 ID AAS49022 standard; DNA; 225 BP.
 XX AC AAS49022;
 XX DT 13-FEB-2002 (first entry)
 XX DE Staphylococcus aureus cellular proliferation inhibitory sequence #246.
 XX KW Antisense; ss; prokaryotic cellular proliferation; antibiotic;
 XX KW antibacterial; drug design.
 XX OS Staphylococcus aureus.
 XX PN WO200170955-A2.
 XX PD 27-SEP-2001.
 XX PF 21-MAR-2001; 2001WO-US0009180.
 XX PR 21-MAR-2000; 2000US-0191078P.
 XX PR 23-MAY-2000; 2000US-0206848P.
 XX PR 26-MAY-2000; 2000US-0207727P.
 XX PR 23-OCT-2000; 2000US-0242578P.
 XX PR 27-NOV-2000; 2000US-0253625P.
 XX PR 22-DEC-2000; 2000US-0257931P.
 XX PR 16-FEB-2001; 2001US-0269308P.
 XX PA (ELIT-) ELITRA PHARM INC.
 XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 XX PI Yamamoto RT, Xu HH;
 XX DR WPI; 2001-6T1135/70.
 XX PT New polynucleotides for the identification and development of
 XX PT antibiotics, comprise sequences of antisense nucleic acids.
 XX PS Claim 1; SEQ ID NO 1599; 511pp; English.
 XX CC The invention relates to antisense inhibitors of genes essential to
 XX CC prokaryotic cellular proliferation, their use in identifying the genes,
 XX CC their use in the discovery of novel antibiotics, the essential genes,
 XX CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
 XX CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
 XX CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
 XX CC useful for the identification of potential new targets for antibiotic
 XX CC development. The antisense nucleic acids can also be used to identify
 XX CC proteins used in proliferation, to express these proteins, and to obtain
 XX CC antibodies capable of binding to the expressed proteins. The proteins can
 XX CC be used to screen compounds in rational drug discovery programmes. The
 XX CC antisense nucleic acid sequence is also useful to screen for homologous

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2004, 21:59:24 ; Search time 636 Seconds
(without alignments)
8803.656 Million cell updates/sec

Title: US-10-712-713-1
Perfect score: 1318
Sequence: 1 atgacacactatcattttgt.....gaaaaatgcgttttaagctt 1318

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1318	100.0	1318	4	Aas00189 S. aureus
2	1314	99.7	1335	9	Add15462 Staphyloc
3	1312.4	99.6	1335	9	Add15464 Staphyloc
4	1311	99.5	1332	7	Acf72770 Staphyloc
5	1310.8	99.5	1314	7	Aca19726 Prokaryot
6	1302.8	98.8	1335	4	Aas54558 Staphyloc
7	1299.8	98.6	1311	4	Aas51660 Staphyloc
8	1293.2	98.1	1351	2	Aav99650 UDP-N-ace
9	1254	95.1	2424	2	Aav74703 Staphyloc
10	922	70.0	1329	6	Abn92833 Staphyloc
11	915.8	69.5	1311	7	Aca47124 Prokaryot
12	633	48.0	660	2	Aav80065 Partial n
13	597.6	45.3	1308	7	Aca22277 Prokaryot
14	570.4	43.3	619	2	Aav53479 DNA encod
15	555	42.1	1371	4	Aas53136 Enterococ
16	553.4	42.0	1335	7	Aca33252 Prokaryot
17	523	39.7	110000	6	Continuation (17 o
18	517.4	39.3	110000	6	Abq67195 Listeria
19	517.4	39.3	110000	6	Continuation (2 of
20	516.4	39.2	1344	7	Aca36934 Prokaryot
21	516.4	39.2	110000	6	Continuation (17 o
22	515	39.1	1332	7	Aca33597 Prokaryot
23	494.8	37.5	738	7	Aca47261 Prokaryot

c	24	456	34.6	110000	6	ABA90521_21
c	25	450.8	34.2	1329	6	ABN68449
c	26	450.8	34.2	110000	6	ABN71527_15
c	27	447.8	34.0	1326	6	ABN68450_15
c	28	447.8	34.0	1329	7	ACA50341
c	29	426.2	32.3	1329	7	ACA48125
c	30	417	31.6	1335	8	ADB11619
c	31	417	31.6	110000	8	ADB12064_14
c	32	402.4	30.5	1335	4	AAS55772_14
c	33	402.4	30.5	1335	7	ACA50019
c	34	400.8	30.4	1332	7	ABX07268
c	35	400.8	30.4	11864	2	AAV52194
c	36	400.8	30.4	110000	7	AB556454_13
c	37	389.6	29.6	888	3	AAA26910
c	38	389.6	29.6	888	4	AAF91592
c	39	389.6	29.6	888	4	AAS08061
c	40	389.6	29.6	888	9	AD67439
c	41	389.6	29.6	888	9	AD673663
c	42	370.2	28.1	1267	2	AAV27381
c	43	370.2	28.1	1267	6	ABQ84849
c	44	370.2	28.1	1267	9	ADC45196
c	45	354.4	26.9	873	9	ADC93633

ALIGNMENTS

RESULT 1
AAS00189
ID AAS00189 standard; DNA; 1318 BP.
XX AC AAS00189;
XX AC AAS00189;

DT 04-JUL-2001 (first entry)

XX S. aureus DNA encoding UDP-N-acetylmuramate:L-alanine ligase, MurC.

XX UDP-N-acetylmuramate:L-alanine ligase; MurC; immunogen; vaccine;
XX antibody; wound infection; cellulitis; burn infection; eyelid infection;
XX food poisoning; joint infection; neonatal conjunctivitis; osteomyelitis;
XX skin infection; scalded skin syndrome; toxic epidermal necrosis;
XX Ritter's disease; Lyell's disease; toxic shock syndrome; endocarditis;
XX ds.

XX Staphylococcus aureus.

XX Key Location/Qualifiers
XX CDS 1..1314
XX /*tag= a
XX /product= "MurC"

XX WO200116292-A2.

XX 08-MAR-2001.

XX 31-AUG-2000; 2000WO-US023773.

XX 01-SEP-1999; 99US-0151933P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Choi GH;

XX WPI; 2001-183259/18.

XX P-PSDB; AAU00828.

XX New isolated nucleic acid for use in diagnosing Staphylococcus infections
XX and in vaccines for eliciting immune responses to the infections.

XX Claim 1; Page 14; 225pp; English.

XX The sequence encodes S. aureus MurC (UDP-N-acetylmuramate:L-alanine
XX ligase). The polynucleotides of the invention are used to detect

CC Staphylococcus nucleic acids in a biological sample from an animal for
CC diagnosing Staphylococcus infections. The polypeptides of the invention
CC are used to detect anti-Staphylococcus antibodies in a biological sample
CC from an animal to diagnose Staphylococcus infections. The polypeptides
CC are also used in vaccines to elicit protective antibodies in an animal to
CC a member of the Staphylococcus genus and for preventing or attenuating an
CC infection caused by a member of the Staphylococcus genus e.g. wound
CC infection, cellulitis, burn infection, eyelid infection, food poisoning,
CC joint infection, neonatal conjunctivitis, osteomyelitis, skin infection,
CC scalded skin syndrome (also known as toxic epidermal necrosis, Ritter's
CC disease and Lyell's disease), toxic shock syndrome and endocarditis. The
CC polynucleotides may also be used in vaccines and for preventing or
CC attenuating a Staphylococcus infection. Antibodies to the polypeptides
CC may be used to purify, detect and target the polypeptides in vitro and
CC in vivo diagnostic and therapeutic methods

XX
SQ Sequence 1318 BP; 464 A; 162 C; 251 G; 441 T; 0 U; 0 Other;

Query Match 100.0%; Score 1318; DB 4; Length 1318;

Best Local Similarity 100.0%; Pred. No. 2.7e-249;

Matches 1318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGACACACTATCATTTTGTGCGAATTAAAGTTCTGGCATGAGTTCAATTAGCACAAATC	60
Db	1	ATGACACACTATCATTTTGTGCGAATTAAAGTTCTGGCATGAGTTCAATTAGCACAAATC	60
QY	61	ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATGAGAACTACGTTATACAGAA	120
Db	61	ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATGAGAACTACGTTATACAGAA	120
QY	121	GTTCCTCTTAGAATAAGGGATAAAAATATACATTTGATGCTATATACAAATACAGAA	180
Db	121	GTTCCTCTTAGAATAAGGGATAAAAATATACATTTGATGCTATATACAAATACAGAA	180
QY	181	GATATGCTAGTTATACAAAGGTAATGCAATTCGCGATGAGCCATGAAGAAATAGTACGTC	240
Db	181	GATATGCTAGTTATACAAAGGTAATGCAATTCGCGATGAGCCATGAAGAAATAGTACGTC	240
QY	241	CATCAATTTGAAATTAGATGTTGTAAGTTATATGATTTTATGAGACATGATTTGATCAA	300
Db	241	CATCAATTTGAAATTAGATGTTGTAAGTTATATGATTTTATGAGACATGATTTGATCAA	300
QY	301	TATACCTTCAGTGTGTAATGTTGTCACATGGTAAAACTTCTCAACAGGTTTATATCA	360
Db	301	TATACCTTCAGTGTGTAATGTTGTCACATGGTAAAACTTCTCAACAGGTTTATATCA	360
QY	361	CATGTTATGATGTTGTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	420
Db	361	CATGTTATGATGTTGTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	420
QY	421	TTGCCTGAAAGTGTATTTTTCGCTTTTTCGCTTTTTCGCTTTTTCGCTTTTTCGCTTTT	480
Db	421	TTGCCTGAAAGTGTATTTTTCGCTTTTTCGCTTTTTCGCTTTTTCGCTTTTTCGCTTTT	480
QY	481	TATAAACCTGATTCGCAATTTATGACAAATATGATTTGATTTGATTTGATTTGATTTAA	540
Db	481	TATAAACCTGATTCGCAATTTATGACAAATATGATTTGATTTGATTTGATTTGATTTAA	540
QY	541	GATATTAATGATGTTTTCGCTTTTTCGCTTTTTCGCTTTTTCGCTTTTTCGCTTTTTCG	600
Db	541	GATATTAATGATGTTTTCGCTTTTTCGCTTTTTCGCTTTTTCGCTTTTTCGCTTTTTCG	600
QY	601	ATTGCTTGGGGTGTATGATGAACTCTACGTAATAATGAAAGTGTGTTTTCGCTTTTTCG	660
Db	601	ATTGCTTGGGGTGTATGATGAACTCTACGTAATAATGAAAGTGTGTTTTCGCTTTTTCG	660
QY	661	TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTTCAAATTTACGGATAAGG	720
Db	661	TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTTCAAATTTACGGATAAGG	720
QY	721	ACTGCTTTTTCGCTTTTTCGCTTTTTCGCTTTTTCGCTTTTTCGCTTTTTCGCTTTTTCG	780
Db	721	ACTGCTTTTTCGCTTTTTCGCTTTTTCGCTTTTTCGCTTTTTCGCTTTTTCGCTTTTTCG	780

QY	781	GCTGACCATACAGTTTAAATGCAATGCTGTAATTTGCGATTAGTTATTTAGAGAAGCTA	840
Db	781	GCTGACCATACAGTTTAAATGCAATGCTGTAATTTGCGATTAGTTATTTAGAGAAGCTA	840
QY	841	GATGTTTACAAATTTAAAGCAATTTAGAAACGTTTGGTGGTGTAAACCTCGTTTCAAT	900
Db	841	GATGTTTACAAATTTAAAGCAATTTAGAAACGTTTGGTGGTGTAAACCTCGTTTCAAT	900
QY	901	GAATCTACAAATTTGCAAAATCAAGTTATTTAGATGATTTGTCACACCATCCAGAGAAAT	960
Db	901	GAATCTACAAATTTGCAAAATCAAGTTATTTAGATGATTTGTCACACCATCCAGAGAAAT	960
QY	961	AGTCTTACAAATTTGAAACAGCAAGAAATATCCACATAAAGAGTTGTTGAGTATTT	1020
Db	961	AGTCTTACAAATTTGAAACAGCAAGAAATATCCACATAAAGAGTTGTTGAGTATTT	1020
QY	1021	CAACACACACTTTCTCTAGAACACAGGCAATTTTAAATGAATTTGCAGAAAGTTTAAAT	1080
Db	1021	CAACACACACTTTCTCTAGAACACAGGCAATTTTAAATGAATTTGCAGAAAGTTTAAAT	1080
QY	1081	AAAGCAGATCGTGTATTTCTTATGTGAAATTTTGGATCAATTTAGAGAAATATCTGGGCA	1140
Db	1081	AAAGCAGATCGTGTATTTCTTATGTGAAATTTTGGATCAATTTAGAGAAATATCTGGGCA	1140
QY	1141	TTAAGCATAAAGATTTAATGATATAAATTTGAAGTGCATCGTTAATTAATGAAGATTTCT	1200
Db	1141	TTAAGCATAAAGATTTAATGATATAAATTTGAAGTGCATCGTTAATTAATGAAGATTTCT	1200
QY	1201	ATTATGATTTAGAACATTTGATTAATGCTGTTATTTTATTTATTTGCGTGCAGGTATTT	1260
Db	1201	ATTATGATTTAGAACATTTGATTAATGCTGTTATTTTATTTATTTGCGTGCAGGTATTT	1260
QY	1261	CAAAATTTACAAATTTGATTAATTTAGATAAATTTAGGATCAAAATCGTTTAAAGCTT	1318
Db	1261	CAAAATTTACAAATTTGATTAATTTAGATAAATTTAGGATCAAAATCGTTTAAAGCTT	1318

RESULT 2

ADD15462

ID ADD15462 standard; DNA; 1335 BP.

XX ADD15462;

AC AC

DT 15-JAN-2004 (first entry)

XX Staphylococcus aureus murC DNA (SeqID 27).

microbial; antimicrobial; membrane biosynthesis; pathogenic;
immunological response; vaccination; surface disinfectant;
personal hygiene application; food preservative; ds; gene; murC;
UDP-N-acetylmuramate-alanine ligase.

OS Staphylococcus aureus.

XX Key Location/Qualifiers

FH CDS 1..1335

FT /start= a

FT /product= "MurC protein"

XX WO2003025007-A2.

XX 27-MAR-2003.

XX 20-SEP-2002; 2002WO-CA001428.

XX 21-SEP-2001; 2001US-0323992P.

XX 21-SEP-2001; 2001US-0324152P.

XX 25-SEP-2001; 2001US-0324692P.

XX 26-OCT-2001; 2001US-0339924P.

XX 29-OCT-2001; 2001US-0350973P.

XX 30-OCT-2001; 2001US-0340924P.

XX 27-NOV-2001; 2001US-0333666P.

```

PR 18-DEC-2001; 2001US-0341732P.
PR 18-DEC-2001; 2001US-0341776P.
PR 19-DEC-2001; 2001US-0341949P.
XX (AFFI-) AFFINIUM PHARM INC.
XX
XX Edwards A, Dharamsi A, Vedadi M, Alam MZ, Awrey D, Beattie B;
PI Domagala M, Houston S, Kanagarajah D, Netherly K, Ng I, Mansoury K;
PI McDonald M, Pinder B, Viola C, Wrezel O;
XX
XX WPI; 2003-468119/44.
DR P-PSDB; ADD15463.
DR
XX
XX Novel crystallized recombinant polypeptides from Staphylococcus aureus,
PT Streptococcus pneumoniae and Escherichia coli and which are involved in
PT membrane biosynthesis, useful as targets for pathogenic bacteria.
XX
XX Claim 20; SEQ ID NO 27; 325pp; English.
XX
XX This invention relates to the structural and functional characterisation
CC of microbial polypeptides from Staphylococcus aureus (S. aureus),
CC Streptococcus pneumoniae (S. pneumoniae) and Escherichia coli (E. coli)
CC that provide novel antimicrobial targets. Specifically, it refers to
CC polypeptides that are involved in membrane biosynthesis, which play a
CC critical role in the life cycle and viability of their pathogenic species
CC of origin, and hence provide valuable drug targets. Furthermore, the
CC invention describes modified version of these proteins that facilitate
CC characterisation by labelling with isotopic or heavy atoms, and also
CC fusion proteins. These proteins provide structural and functional
CC information to aid the discovery of therapeutic molecules to treat
CC disorders associated with a particular pathogenic species. As such, they
CC are useful for inducing an immunological response in an individual and as
CC an antigen for vaccination purposes. The polypeptides are also useful for
CC developing antimicrobial agents for use as surface disinfectants,
CC personal hygiene applications and as food preservatives or in treating
CC food products to eliminate potential pathogens. This polynucleotide
CC sequence is DNA predicted from the genomic sequence of S. aureus UDP-N-
CC acetylmuramate-alanine ligase (murC) of the invention.
XX
XX Sequence 1335 BP; 472 A; 161 C; 255 G; 447 T; 0 U; 0 Other;
XX
XX Query March 99.7%; Score 1314; DB 9; Length 1335;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-248;
XX Matches 1314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 ATGACACACTATCATTTTGTGCGAATTAAAGGTTCTGGCATGAGTTTCATTAGCACAAATC 60
XX |||||
XX 22 ATGACACACTATCATTTTGTGCGAATTAAAGGTTCTGGCATGAGTTTCATTAGCACAAATC 81
XX |||||
XX
XX 61 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTTATTACAGAA 120
XX |||||
XX 82 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTTATTACAGAA 141
XX |||||
XX
XX 121 GTTGCTCTTAGAATAAGGGATAAAATATATACATTTGATGCTTAATAACATAAAAGAA 180
XX |||||
XX 142 GTTGCTCTTAGAATAAGGGATAAAATATATACATTTGATGCTTAATAACATAAAAGAA 201
XX |||||
XX
XX 181 GATATGTTAGTTATACAGGTAAGTCAATTCGGATAGCCATGAGAAATAGTAGTGCA 240
XX |||||
XX 202 GATATGTTAGTTATACAGGTAAGTCAATTCGGATAGCCATGAGAAATAGTAGTGCA 261
XX |||||
XX
XX 241 CATCAATTTAGAAATTTAGATGTTTGAAGTTATATGATTTTATAGGACAGATTATTCATCA 300
XX |||||
XX 262 CATCAATTTAGAAATTTAGATGTTTGAAGTTATATGATTTTATAGGACAGATTATTCATCA 321
XX |||||
XX
XX 301 TATACCTTCAGTGTAACTGGTGCAATGCTGAAACCTTCTACACAGGTTTATATCA 360
XX |||||
XX 322 TATACCTTCAGTGTAACTGGTGCAATGCTGAAACCTTCTACACAGGTTTATATCA 381
XX |||||
XX
XX 361 CATGTTATGAATGGTGATAAAAGACTTCATTTTAAATGGTGATGGCACAGGTATGGGA 420
XX |||||
XX 382 CATGTTATGAATGGTGATAAAAGACTTCATTTTAAATGGTGATGGCACAGGTATGGGA 441
XX |||||
XX

```

RESULT 3

ADD15464

ID ADD15464 standard; DNA; 1335 BP.

XX ADD15464;

XX AC

DT 15-JAN-2004 (first entry)

DE Staphylococcus aureus murC DNA (SeqID 29).

KW microbial; antimicrobial; membrane biosynthesis; pathogenic;
 immunological response; vaccination; surface disinfectant;

KW personal hygiene application; food preservative; ds; gene; murc;
KW UDP-N-acetylmuramate-alanine ligase.
XX

OS Staphylococcus aureus.

XX Key Location/Qualifiers

XX CDS 1..1335

XX /tag= a

XX /product= "MurC protein"

XX WO2003025007-A2.

XX

XX 27-MAR-2003.

XX 20-SEP-2002; 2002WO-CA001428.

XX 21-SEP-2001; 2001US-0323992P.

XX 21-SEP-2001; 2001US-0324152P.

XX 25-SEP-2001; 2001US-0324692P.

XX 26-OCT-2001; 2001US-0339924P.

XX 29-OCT-2001; 2001US-0350973P.

XX 30-OCT-2001; 2001US-0340924P.

XX 27-NOV-2001; 2001US-0333666P.

XX 18-DEC-2001; 2001US-0341732P.

XX 18-DEC-2001; 2001US-0341776P.

XX 19-DEC-2001; 2001US-0341949P.

XX (AFFI-) AFFINIUM PHARM. INC.

XX

XX Edwards A, Vedadi M, Alam MZ, Awrey D, Beattie B;

XX Domagala M, Houston S, Kanagarajah D, Nethery K, Ng I, Mansoury K;

XX McDonald M, Pinder B, Viola C, Wrezel O;

XX WPI; 2003-468119/44.

XX P-PSDB; ADD15465.

XX Novel crystallized recombinant polypeptides from Staphylococcus aureus,

XX Streptococcus pneumoniae and Escherichia coli and which are involved in

XX membrane biosynthesis, useful as targets for pathogenic bacteria.

XX Claim 20; SEQ ID NO 29; 325pp; English.

XX This invention relates to the structural and functional characterisation

XX of microbial polypeptides from Staphylococcus aureus (S. aureus),

XX Streptococcus pneumoniae (S. pneumoniae) and Escherichia coli (E. coli)

XX that provide novel antimicrobial targets. Specifically, it refers to

XX polypeptides that are involved in membrane biosynthesis, which play a

XX critical role in the life cycle and viability of their pathogenic species

XX of origin, and hence provide valuable drug targets. Furthermore, the

XX invention describes modified version of these proteins that facilitate

XX fusion proteins. These proteins provide structural and functional

XX information to aid the discovery of therapeutic molecules to treat

XX disorders associated with a particular pathogenic species. As such, they

XX are useful for inducing an immunological response in an individual and as

XX an antigen for vaccination purposes. The polypeptides are also useful for

XX developing antimicrobial agents for use as surface disinfectants,

XX personal hygiene applications and as food preservatives or in treating

XX food products to eliminate potential pathogens. This polynucleotide

XX sequence is the experimentally predicted DNA of S. aureus UDP-N-

XX acetylmuramate-alanine ligase (murC) of the invention.

XX

XX SQ Sequence 1335 BP; 472 A; 160 C; 255 G; 448 T; 0 U; 0 Other;

XX

XX Query Match 99.6%; Score 1312.4; DB 9; Length 1335;

XX Best Local Similarity 99.9%; Pred. No. 3.4e-248;

XX Matches 1313; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX

XX 1 ATGACACACTATCATTTTGGCGAATTAAAGGTTCTGGCATGAGTTTCATTAGCACAAATC 60

XX

XX 22 ATGACACACTATCATTTTGGCGAATTAAAGGTTCTGGCATGAGTTTCATTAGCACAAATC 81

XX

XX 61 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATGAGAACTACGTATTACAGAA 120

XX

XX

XX

XX

XX

XX

Db 82 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTACAGAA 141
Qy 121 GTTGCTCTAGAAATAAGGGGATAAAAATATTACCATTTTGATGCTTAATAACATAAAGAA 180
Db 142 GTTGCTCTAGAAATAAGGGGATAAAAATATTACCATTTTGATGCTTAATAACATAAAGAA 201
Qy 181 GATATGGTAGTTATACAAAGGTAATGCAATTCGGAGTAGCCATGAAGAAATAGTACGTGA 240
Db 202 GATATGGTAGTTATACAAAGGTAATGCAATTCGGAGTAGCCATGAAGAAATAGTACGTGA 261
Qy 241 CATCAATTTGAATTTAGATGTTTAAAGTTATATAATGATTTTATAGGACAGATTTATGATCA 300
Db 262 CATCAATTTGAATTTAGATGTTTAAAGTTATATAATGATTTTATAGGACAGATTTATGATCA 321
Qy 301 TATACCTTCAGTAGCTGTAACTGGTCGACATGCTGTAATAAACTTCTACACACAGTTTATATCA 360
Db 322 TATACCTTCAGTAGCTGTAACTGGTCGACATGCTGTAATAAACTTCTACACACAGTTTATATCA 381
Qy 361 CATGTTATGAATGGTGATATAAAAGACTTCATTTTTAAATTTGGTGATGGCAAGGTATGGGA 420
Db 382 CATGTTATGAATGGTGATATAAAAGACTTCATTTTTAAATTTGGTGATGGCAAGGTATGGGA 441
Qy 421 TTGCTGTAAGTGAATTTTGGTGATGCAATTCGCTTTTGGAGCATGCTGATATAGAGTCACTTTTAAAGT 480
Db 442 TTGCTGTAAGTGAATTTTGGTGATGCAATTCGCTTTTGGAGCATGCTGATATAGAGTCACTTTTAAAGT 501
Qy 481 TATAAACCCTGATTAACGCAATTAATGACAAATATTGATTTTCGATCATCTCTGATTTATTTAAA 540
Db 502 TATAAACCCTGATTAACGCAATTAATGACAAATATTGATTTTCGATCATCTCTGATTTATTTAAA 561
Qy 541 GATATTAATGAATGGTGATGCAATTCGCTTTTGGAGCATGCTGATATAGAGTCACTTTTAAAGT 600
Db 562 GATATTAATGAATGGTGATGCAATTCGCTTTTGGAGCATGCTGATATAGAGTCACTTTTAAAGT 621
Qy 601 ATTGCTTTGGTGATGATGAACATCTACGTAATAAATTTGAAGCATGTTTCCATTTATTTAT 660
Db 622 ATTGCTTTGGTGATGATGAACATCTACGTAATAAATTTGAAGCATGTTTCCATTTATTTAT 681
Qy 661 TATGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTTCAAATTTACGGATAAAGT 720
Db 682 TATGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTTCAAATTTACGGATAAAGT 741
Qy 721 ACTGCTTTTGGTGATGATGATGGTGGTGATGATGATGATGATGATGATGATGATGATGAT 780
Db 742 ACTGCTTTTGGTGATGATGATGGTGGTGATGATGATGATGATGATGATGATGATGATGAT 801
Qy 781 GGTGACCATACAGTTTAAATGCAATTCGCTGATGATGATGATGATGATGATGATGATGATGAT 840
Db 802 GGTGACCATACAGTTTAAATGCAATTCGCTGATGATGATGATGATGATGATGATGATGATGAT 861
Qy 841 GATGTTACAAATATTAAAGAGACATTGAAAGCTTTGGTGTTTAAAGCTGTTTCAAT 900
Db 862 GATGTTACAAATATTAAAGAGACATTGAAAGCTTTGGTGTTTAAAGCTGTTTCAAT 921
Qy 901 GAAACTACAAATGCAAAATCAAGTTTATCTAGATGATGATGATGATGATGATGATGATGATGAT 960
Db 922 GAAACTACAAATGCAAAATCAAGTTTATCTAGATGATGATGATGATGATGATGATGATGATGAT 981
Qy 961 AGTGCTCAATGTAAGAACAGCAGAAAGAAATATCCACATAAAGAGTTGTTGTCAGTATTT 1020
Db 982 AGTGCTCAATGTAAGAACAGCAGAAAGAAATATCCACATAAAGAGTTGTTGTCAGTATTT 1041
Qy 1021 CAACACACACTTTCTCTAGAACACAGGCATTTTAAATGATTTTGCAGAAAGTTTAACT 1080
Db 1042 CAACACACACTTTCTCTAGAACACAGGCATTTTAAATGATTTTGCAGAAAGTTTAACT 1101
Qy 1081 AAACAGATCGTGATTTCTTATGTAATTTTGGATCAATTTAGAGAAAATACTCGCGCA 1140
Db 1102 AAACAGATCGTGATTTCTTATGTAATTTTGGATCAATTTAGAGAAAATACTCGCGCA 1161
Qy 1141 TTAACGATACAGATTTTAAATGATAAAATTTGAAGTGCATCGTATTAATTAATGAAGATTCT 1200

Db 1222 ATTAATGTTAGAACAAATTTGATAATGCTGTTATTTATTTATTTATGCTGCGGTGATATT 1281
Qy 1261 CAAAATTTACAAATGCAATTTAGATAATTTAGCATGAAATATCGGTTT 1311
Db 1282 CAAAATTTACAAATGCAATTTAGATAATTTAGCATGAAATATCGGTTT 1332

RESULT 5

ACAL9726
ID ACAL9726 standard; DNA; 1314 BP.

XX AC
XX AC
XX ACAL9726;

DT 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #1383.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.

XX Staphylococcus aureus.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851P.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlson KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR P-PSDB; ABUI5856.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 7596; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1314 BP; 464 A; 161 C; 250 G; 439 T; 0 U; 0 Other;

Query Match 99.5%; Score 1310.8; DB 7; Length 1314;
Best Local Similarity 99.8%; Pred. No. 7e-248;
Matches 1312; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGACACACTATCATTTTGTTCGGAATTAAGGTTCTGGCATGAGTTCAATTAGCACAATC 60
Db 1 ATGACACACTATCATTTTGTTCGGAATTAAGGTTCTGGCATGAGTTCAATTAGCACAATC 60
Qy 61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTTATTACAGAA 120
Db 61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTTATTACAGAA 120
Qy 121 GTTGCTCTTAGAAATAAGGGGATAAAATATTACCAATTTGATCTAATACATAAAGAA 180
Db 121 GTTGCTCTTAGAAATAAGGGGATAAAATATTACCAATTTGATCTAATACATAAAGAA 180
Qy 181 GATATGTTAGTTATATCAAGGTAATGCAATTCGGGAGTAGCCATGAAGAAATAGTACGTGCA 240
Db 181 GATATGTTAGTTATATCAAGGTAATGCAATTCGGGAGTAGCCATGAAGAAATAGTACGTGCA 240
Qy 241 CATCAATTGAAATAGATGTTGTAAGTTAATGATTTTTTTAGGACAGATTTATTGATCAA 300
Db 241 CATCAATTGAAATAGATGTTGTAAGTTAATGATTTTTTTAGGACAGATTTATTGATCAA 300
Qy 301 TATATCTCAGTAGCTGTAACTGGTGCAACATGCTGTAAGAACTTCTACACAGGTTATTATCA 360
Db 301 TATATCTCAGTAGCTGTAACTGGTGCAACATGCTGTAAGAACTTCTACACAGGTTATTATCA 360
Qy 361 CATGTTATGAATGGTGATATAAAGAGCTTCAATTTTAAATTTGGTGCATGGCAGGTTATGGGA 420
Db 361 CATGTTATGAATGGTGATATAAAGAGCTTCAATTTTAAATTTGGTGCATGGCAGGTTATGGGA 420
Qy 421 TTGCTGCTGAAAGTGATTTTTCGCTTTTGGAGGCATGTGAATATAGAGCTCACTTTTAAAGT 480
Db 421 TTGCTGCTGAAAGTGATTTTTCGCTTTTGGAGGCATGTGAATATAGAGCTCACTTTTAAAGT 480
Qy 481 TATAAACCTGATTACGCAATTTATGCAAAATATGATTTTCGATCATCTCTGATTTTAAAGT 540
Db 481 TATAAACCTGATTACGCAATTTATGCAAAATATGATTTTCGATCATCTCTGATTTTAAAGT 540
Qy 541 GATATTAATGATTTTTCGCTTTTTCGCTTTTGGAGGCATGTGAATATAGAGCTCACTTTTAAAGT 600
Db 541 GATATTAATGATTTTTCGCTTTTTCGCTTTTGGAGGCATGTGAATATAGAGCTCACTTTTAAAGT 600
Qy 601 ATTGCTGGGGTGATGATGAACATCTACGTAATAATTTGAAGCAGATGTTTCAATTTATTAT 660
Db 601 ATTGCTGGGGTGATGATGAACATCTACGTAATAATTTGAAGCAGATGTTTCAATTTATTAT 660
Qy 661 TATGATTTAAAGATTTCGGATGACATTTATGCTCAAAATATTCAAAATTTACGGAATAAGGT 720
Db 661 TATGATTTAAAGATTTCGGATGACATTTATGCTCAAAATATTCAAAATTTACGGAATAAGGT 720
Qy 721 ACTGCTTTTGTATGTTATGCTGATGTTTATGATCACTTCTCTCCCAATAT 780
Db 721 ACTGCTTTTGTATGTTATGCTGATGTTTATGATCACTTCTCTCCCAATAT 780
Qy 781 GGTGACCACATACAGTTTAAATTCGATTTAGTATTTAGAGAGCTA 840
Db 781 GGTGACCACATACAGTTTAAATTCGATTTAGTATTTAGAGAGCTA 840
Qy 841 GATGTTCAAAATATAAGAGAGCTTAAAGAGCTTTGGTGGTGTAAAGCTGTTTCAAT 900
Db 841 GATGTTCAAAATATAAGAGAGCTTAAAGAGCTTTGGTGGTGTAAAGCTGTTTCAAT 900

QY 901 GAAACTACAAATGCAATCAAGTTATTGTAGATGATTAATGACACACCATCCCAAGAAATTT 960
Db |||||
QY 901 GAAACTACAAATGCAATCAAGTTATTGTAGATGATTAATGACACACCATCCCAAGAAATTT 960
Db |||||
QY 961 AGTGCTACAAATGAAACAGCAGCAAGAAATATCCACATAAAGAGTTGTCAGTATTT 1020
Db |||||
QY 961 AGTGCTACAAATGAAACAGCAGCAAGAAATATCCACATAAAGAGTTGTCAGTATTT 1020
Db |||||
QY 1021 CAACCCACACACTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAGTTTAAAGT 1080
Db |||||
QY 1021 CAACCCACACACTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAGTTTAAAGT 1080
Db |||||
QY 1081 AAGAGCAGATCGTGATTTCTTATGTGAATTTTGGATCAATTAGAGAAATATCTGGCGCA 1140
Db |||||
QY 1081 AAGAGCAGATCGTGATTTCTTATGTGAATTTTGGATCAATTAGAGAAATATCTGGCGCA 1140
Db |||||
QY 1141 TTAACGATACAAAGTTTAAATGAATTTGAAGTGATGCTGTTAAATTAATGAATTTCT 1200
Db |||||
QY 1141 TTAACGATACAAAGTTTAAATGAATTTGAAGTGATGCTGTTAAATTAATGAATTTCT 1200
Db |||||
QY 1201 ATTAATGTATTAGAACCAATTTGATAATGCTGTTATTTTATTTATGCGTGAGGTGATATT 1260
Db |||||
QY 1201 ATTAATGTATTAGAACCAATTTGATAATGCTGTTATTTTATTTATGCGTGAGGTGATATT 1260
Db |||||
QY 1261 CAAAAATTACAAATGATATTTAGATAAAATTAGGCATGAAAAATGCGTTTTAA 1314
Db |||||
QY 1261 CAAAAATTACAAATGATATTTAGATAAAATTAGGCATGAAAAATGCGTTTTAA 1314
Db |||||

RESULT 6

AAS54558

ID AAS54558 standard; DNA; 1335 BP.

XX

AC AAS54558;

XX

13-FEB-2002 (first entry)

XX

Staphylococcus aureus DNA for cellular proliferation protein #870.

XX

Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;

KW

antibacterial; drug design.

KW

XX

Staphylococcus aureus.

OS

WO200170955-A2.

XX

27-SEP-2001.

PD

21-MAR-2001; 2001WO-US009180.

XX

21-MAR-2000; 2000US-0191078P.

PR

23-MAY-2000; 2000US-0206848P.

PR

26-MAY-2000; 2000US-0207272P.

PR

23-OCT-2000; 2000US-0242578P.

PR

27-NOV-2000; 2000US-0253625P.

PR

22-DEC-2000; 2000US-0257931P.

PR

16-FEB-2001; 2001US-0289308P.

XX

(ELIT-) ELITRA PHARM INC.

XX

Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI

Yamamoto-RT, Xu HH;

PI

WPI; 2001-611495/70.

DR

P-PSDB; AAU36699.

XX

New polynucleotides for the identification and development of

PT

antibiotics, comprise sequences of antisense nucleic acids.

XX

Claim 27; SEQ ID NO 8195; 511pp; English.

XX

The invention relates to antisense inhibitors of genes essential to

CC

prokaryotic cellular proliferation, their use in identifying the genes,

CC

CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 1335 BP; 472 A; 164 C; 255 G; 444 T; 0 U; 0 Other;

Query Match 98.8%; Score 1302.8; DB 4; Length 1335;
Best Local Similarity 99.5%; Pred. No. 2.6e-246;
Matches 1307; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGACACACTATCATTTTGTGGAAATTAAGGTTCTGGCAATGAGTTTCATTAGCACAATC 60
Db |||||
QY 22 ATGACACACTATCATTTTGTGGAAATTAAGGTTCTGGCAATGAGTTTCATTAGCACAATC 81
Db |||||
QY 61 ATGCATGATTAGGACATGAAGTTCAAGGATCCGATATTCAGACTACGTTATTACAGAA 120
Db |||||
QY 82 ATGCATGATTAGGACATGAAGTTCAAGGATCCGATATTCAGACTACGTTATTACAGAA 141
Db |||||
QY 121 GTTGCTCTTAGAAATAAGGGGATAAAATATTAACCATTTTCATGCTAATAACATAAGAA 180
Db |||||
QY 142 GTTGCTCTTAGAAATAAGGGGATAAAATATTAACCATTTTCATGCTAATAACATAAGAA 201
Db |||||
QY 181 GATATGGTAGTTATACAAGGTAATGCATTGCGAGTAGCCATGAAGAAATAGTAGTGCA 240
Db |||||
QY 202 GATATGGTAGTTATACAAGGTAATGCATTGCGAGTAGCCATGAAGAAATAGTAGTGCA 261
Db |||||
QY 241 CATCAATTCGAAATAGATGTTGTAAGTTAATGATTTTATGATTTTATGATTTATGATCA 300
Db |||||
QY 262 CATCAATTCGAAATAGATGTTGTAAGTTAATGATTTTATGATTTTATGATTTATGATCA 321
Db |||||
QY 301 TATACCTTCAGTAGCTTAACCTGGTGACATGTAAGTAACTCTCAACACAGGTTTATATCA 360
Db |||||
QY 322 TATACCTTCAGTAGCTTAACCTGGTGACATGTAAGTAACTCTCAACACAGGTTTATATCA 381
Db |||||
QY 361 CATGTTATCAATGGTGATATAAGAACTTCATTTTAAATTTGGTGATGSCACAGGTATCGGA 420
Db |||||
QY 382 CATGTTATCAATGGTGATATAAGAACTTCATTTTAAATTTGGTGATGSCACAGGTATCGGA 441
Db |||||
QY 421 TTGCTCGAAAGTGATTTATTTGCTTTTGGGCAATGTAATATAGACGTCATTTTAAAT 480
Db |||||
QY 442 TTGCTCGAAAGTGATTTATTTGCTTTTGGGCAATGTAATATAGACGTCATTTTAAAT 501
Db |||||
QY 481 TATAACCTGATTCAGCAATATGACAAATATGATTTGATTTGATTCATCTCGATTTTAA 540
Db |||||
QY 502 TATAACCTGATTCAGCAATATGACAAATATGATTTGATTTGATTCATCTCGATTTTAA 561
Db |||||
QY 541 GATATTAAATGATGTTTTCATGTCATTTCCAGAAATGCGACATATGTTTAAAGAGGTATT 600
Db |||||
QY 562 GATATTAAATGATGTTTTCATGTCATTTCCAGAAATGCGACATATGTTTAAAGAGGTATT 621
Db |||||
QY 601 ATTGCTTGGGGTGATGATGAACATCTACGTAATTTGAAGCAGATGTTTCAATTTATAT 660
Db |||||
QY 622 ATTGCTTGGGGTGATGATGAACATCTACGTAATTTGAAGCAGATGTTTCAATTTATAC 681
Db |||||
QY 661 TATGATTTAAAGATTCGGATGATGATTTATGCTCAAAATATTCAAATACGATAAGGT 720
Db |||||
QY 682 TATGATTTAAAGATTCGGATGATGATTTATGCTCAAAATATTCAAATACGATAAGGT 741
Db |||||
QY 721 ACTGCTTTTATGATGATGATGATGATGATTTTATGATGATGATTTTATGATGATTTTAT 780
Db |||||

Db 742 ACTGCTTTGATGTGTATGGATGGTGTATGATCATCTCTGCTCCACATAT 801
QY 781 GGTGACCATACAGTTTTAAATGCAATAGCTGTAAATTCGATTAATTTAGAGAGCTA 840
Db 802 GGTGACCATACAGTTTTAAATGCAATAGCTGTAAATTCGATTAATTTAGAGAGCTA 861
QY 841 GATGTTACAAATATTAAGAGCAATTAGAAACGTTTGGTGGTGTAAACGTCGTTCAAT 900
Db 862 GATGTTACAAATATTAAGAGCAATTAGAAACGTTTGGTGGTGTAAACGTCGTTCAAT 921
QY 901 GAAACTACAAATTCGAAATCAAGTTATGTTAGATGATTATGACACCATCCAGAGAAAT 960
Db 922 GAAACTACAAATTCGAAATCAAGTTATGTTAGATGATTATGACACCATCCAGAGAAAT 981
QY 961 AGTGCTACAAATGAAACAGACAGCAAGAAATATCCACATAAAGAGTTGTTGCAGTATTT 1020
Db 982 AGTGCTACAAATGAAACAGACAGCAAGAAATATCCACATAAAGAGTTGTTGCAGTATTT 1041
QY 1021 CAACACACATCTCTAGAACACAGGCAATTTTAAATGAATTTGCAGAAAGTTTAAAT 1080
Db 1042 CAACACACATCTCTAGAACACAGGCAATTTTAAATGAATTTGCAGAAAGTTTAAAT 1101
QY 1081 AAGCAGATCGTGATCTTATGCTGAAATTTTGGATCAATTAGAGAAATATCTGGCGCA 1140
Db 1102 AAGCAGATCGTGATCTTATGCTGAAATTTTGGATCAATTAGAGAAATATCTGGCGCA 1161
QY 1141 TTAACGATACAAAGATTTAATGATAAAATGAAGGTGCAATCGTTAATTAATGAAGATCT 1200
Db 1162 TTAACGATACAAAGATTTAATGATAAAATGAAGGTGCAATCGTTAATTAATGAAGATCT 1221
QY 1201 ATTAATGATTAAGAACAAATTTGATTAATGCTGTTATTTATTTATGGTGCAAGTGATTT 1260
Db 1222 ATTAATGATTAAGAACAAATTTGATTAATGCTGTTATTTATTTATGGTGCAAGTGATTT 1281
QY 1261 CAAAATTAACAAATCATATTAGATAAATTAAGGATGCAAGAAATCGGTTTAA 1314
Db 1282 CAAAATTAACAAATCATATTAGATAAATTAAGGATGCAAGAAATCGGTTTAA 1335

RESULT 7

AAS51660

ID AAS51660 standard; DNA; 1311 BP.

XX AC

AAS51660;

XX DT

13-FEB-2002 (first entry)

XX DE

Staphylococcus aureus DNA for cellular proliferation protein #77.

XX KW

Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;

XX KW

antibacterial; drug design.

XX OS

Staphylococcus aureus.

XX XX

WO200170955-A2.

XX PN

27-SEP-2001.

XX PF

21-MAR-2001; 2001WO-US009180.

XX PR

21-MAR-2000; 2000US-0191078P.

XX PR

23-MAY-2000; 2000US-0206848P.

XX PR

26-MAY-2000; 2000US-0207727P.

XX PR

23-OCT-2000; 2000US-0242578P.

XX PR

27-NOV-2000; 2000US-0253625P.

XX PR

22-DEC-2000; 2000US-0257931P.

XX PR

16-FEB-2001; 2001US-0269308P.

XX PA

(ELIT-) ELITRA PHARM INC.

XX XX

Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

WPI; 2001-611495/70.
P-PSDB; AAU33801.New polynucleotides for the identification and development of
antibiotics, comprise sequences of antisense nucleic acids.

Claim 27; SEQ ID NO 4242; 511pp; English.

The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence encodes an essential prokaryotic cellular proliferation protein. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 1311 BP; 462 A; 164 C; 250 G; 435 T; 0 U; 0 Other;

Query Match 98.6%; Score 1299.8; DB 4; Length 1311;

Best Local Similarity 99.5%; Pred. No. 1e-245;

Matches 1304; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGCACACATATCATTTTGTGCGAATTAAGGTTCTGCGATGAGTTCATTAGCACAAATC 60

Db 1 ATGCACACATATCATTTTGTGCGAATTAAGGTTCTGCGATGAGTTCATTAGCACAAATC 60

QY 61 ATGCATGATTTAGGACATGAAAGTTCAAGGATCGGATATTTGAGAACTACGTATTTACAGAA 120

Db 61 ATGCATGATTTAGGACATGAAAGTTCAAGGATCGGATATTTGAGAACTACGTATTTACAGAA 120

QY 121 GTTGCTCTTAGAAATAAGGGGATATAAATATATACATTTGATGCTTAATAACATAAAGAA 180

Db 121 GTTGCTCTTAGAAATAAGGGGATATAAATATATACATTTGATGCTTAATAACATAAAGAA 180

QY 181 GATATGCTAGTTATACAGGTAATGCTTCGCGAGTAGCATGGAAGAATAGTACGTGCA 240

Db 181 GATATGCTAGTTATACAGGTAATGCTTCGCGAGTAGCATGGAAGAATAGTACGTGCA 240

QY 241 CATCAATTTGAAATTAGATGTTGTAAGTTATATGATTTTTTTAGGACAGATTTATGATCAA 300

Db 241 CATCAATTTGAAATTAGATGTTGTAAGTTATATGATTTTTTTAGGACAGATTTATGATCAA 300

QY 301 TATACCTCAGTAGCTGTAACTGGTGCCATCGGTAAACCTCTCAACAGGTTTATATCA 360

Db 301 TATACCTCAGTAGCTGTAACTGGTGCCATCGGTAAACCTCTCTACACAGGTTTATATCA 360

QY 361 CATGTTATGATGCTGATATAAAGCACTTCATTTTTTAATTTGGTGAGGACAGGATGGGA 420

Db 361 CATGTTATGATGCTGATATAAAGCACTTCATTTTTTAATTTGGTGAGGACAGGATGGGA 420

QY 421 TTGCTGAAAGTGATTTATTTTCGCTTTTGGAGCATGTGAATATAGACGTCACATTTTAAAT 480

Db 421 TTGCTGAAAGTGATTTATTTTCGCTTTTGGAGCATGTGAATATAGACGTCACATTTTAAAT 480

QY 481 TATAAACCTGATACGCAATTTATGCAAAATATGATTTTCGATCATCTCGATATTTTAAA 540

Db 481 TATAAACCTGATACGCAATTTATGCAAAATATGATTTTCGATCATCTCGATATTTTAAA 540

QY 541 GATATTAATGATGTTTGTGATGATTTCCAGAAATGCGACATATGTTTAAAAAGGATTT 600

Db 541 GATATTAATGATGTTTGTGATGATTTCCAGAAATGCGACATATGTTTAAAAAGGATTT 600

QY 601 ATTGCTGGGTCATGATGAAATCTACGTAAATTTGAACGAGATGTTCCAAATTTATTAT 660
Db 601 ATTGCTGGGTCATGATGAAATCTACGTAAATTTGAACGAGATGTTCCAAATTTATTAT 660
QY 661 TATGGATTTAAAGATTCGATGACATTTATGCTCAAAATATTCAAATTCAGGATAAAGT 720
Db 661 TATGGATTTAAAGATTCGATGACATTTATGCTCAAAATATTCAAATTCAGGATAAAGT 720
QY 721 ACTGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 721 ACTGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 GGTGACCATACAGTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 781 GGTGACCATACAGTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 GATGTTACAAATTTAAAGAGCATTTAGAACCTTTGGTGTAAACCTCGTTTCAAT 900
Db 841 GATGTTACAAATTTAAAGAGCATTTAGAACCTTTGGTGTAAACCTCGTTTCAAT 900
QY 901 GAACTACAAATGCAATCAAGTTTATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 901 GAACTACAAATGCAATCAAGTTTATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 AGTGCTACAAATGAAACAGCAGCAAGAAATATCCACATAAAGAGTTGTCAGATATT 1020
Db 961 AGTGCTACAAATGAAACAGCAGCAAGAAATATCCACATAAAGAGTTGTCAGATATT 1020
QY 1021 CAACACACACTTCTCTAGAACACAGCAATTTTAAATGCAATTTGAGAAAGTTTAAAT 1080
Db 1021 CAACACACACTTCTCTAGAACACAGCAATTTTAAATGCAATTTGAGAAAGTTTAAAT 1080
QY 1081 AAGCAGATCGTATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db 1081 AAGCAGATCGTATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 TTAACGATACAGATTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Db 1141 TTAACGATACAGATTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 ATTAATGATTAAGACAAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db 1201 ATTAATGATTAAGACAAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1261 CAATAATTTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1311
Db 1261 CAATAATTTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1311

RESULT 8
AAV99650
ID AAV99650 standard; DNA; 1351 BP.
XX
AC AAV99650;
DT 17-MAR-1999 (first entry)
XX
DE UDP-N-acetylmuramate:L-alanine ligase (MurC polypeptide) encoding DNA.
KW MurC gene; UDP-N-acetylmuramate:L-alanine ligase; MurC polypeptide;
KW bacterial; infection; H. pylori; cancer; ulcer; gastritis; vaccine;
KW immunogen; drug; genetic immunisation; ds.
XX Staphylococcus aureus.
XX
FH Location/Qualifiers
FT 22..1335
FT /tag= a
FT /gene= "MurC"
FT /product= "UDP-N-acetylmuramate:L-alanine ligase (MurC
FT polypeptide)"
XX
PN EP889123-A2.

XX 07-JAN-1999.
XX 26-JUN-1998; 98EP-00305064.
XX 03-JUL-1997; 97US-0052720P.
XX (SMIK) SMITHLINE BEECHAM CORP.
XX (SMIK) SMITHLINE BEECHAM PLC.
XX Wallis NG, Burnham MKR;
XX WPI; 1999-062655/06.
XX P-PSDB; AAW87771.
XX
XX New isolated MurC polypeptide from Staphylococcus aureus and related
XX nucleic acid - useful in diagnosis, treatment and prevention of bacterial
XX infections.
XX
XX Claim 2; Page 3-4; 39pp; English.
XX
XX The present sequence represents a MurC gene encoding a Staphylococcus
XX aureus UDP-N-acetylmuramate:L-alanine ligase (MurC polypeptide). Host
XX cell containing an expression system comprising the MurC gene can be used
XX for the recombinant production of the polypeptide. Agonists or the MurC
XX polypeptide are used to treat conditions requiring increased activity or
XX expression of the polypeptide. Antagonists, inhibitory nucleic acid or
XX competitive polypeptide are useful for inhibiting the polypeptide e.g.
XX bacterial (especially S. aureus) infections. They are also useful against
XX Helicobacter pylori infections and related cancers, ulcers and gastritis.
XX The antibacterial agents are useful to treat in-dwelling devices for
XX infection prevention or generally as wound treatments to prevent adhesion
XX of bacteria to matrix proteins. The MurC polypeptide is also useful for
XX diagnosing or prognosing a (susceptibility to) disease, for raising
XX antibodies; to identify modulators or specific receptors; in rational
XX drug design and as an immunogen for vaccines. The MurC gene sequences are
XX useful in antisense/ribozyme therapeutics; to detect mutant MurC gene;
XX for chromosomal mapping; to determine bacterial serotype; and for genetic
XX immunisation
XX
XX Sequence 1351 BP; 471 A; 166 C; 260 G; 454 T; 0 U; 0 Other;

Query Match 98.1%; Score 1293.2; DB 2; Length 1351;
Best Local Similarity 99.0%; Pred. No. 244;
Matches 1301; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1 ATGCACACTATCATTTTGTGCGAATTAAGGTTCTGGCATGAGTTCAATAGCACAATC 60
Db 22 ATGCACACTATCATTTTGTGCGAATTAAGGTTCTGGCATGAGTTCAATAGCACAATC 81
QY 61 ATGCATGATTTAGGACATGAGTTTCAAGGATCGGATATTGAGAACTACGTTATTACAGAA 120
Db 82 ATGCATGATTTAGGACATGAGTTTCAAGGATCGGATATTGAGAACTACGTTATTACAGAA 141
QY 121 GTTGCTCTTAGAATAAAGGGGATAAAATATTACCATTTGATGCTTAATACATAAAGAA 180
Db 142 GTTGCTCTTAGAATAAAGGGGATAAAATATTACCATTTGATGCTTAATACATAAAGAA 201
QY 181 GATATGCTGATTTATACAGGTAATGCTTCGCGATAGCCATGAGCAATAGTACGTGCA 240
Db 202 GATATGCTGATTTATACAGGTAATGCTTCGCGATAGCCATGAGCAATAGTACGTGCA 261
QY 241 CATCAATTTGAAATTTAGATGTTGTTAGTTATATGATTTTATAGGACAGATTTATGATCAA 300
Db 262 CATCAATTTGAAATTTAGATGTTGTTAGTTATATGATTTTATAGGACAGATTTATGATCAA 321
QY 301 TATACCTTCAGTAGCTGTAACTGGTGCCATGATGTAATCTTCTACACAGGTTTATTATCA 360
Db 322 TATACCTTCAGTAGCTGTAACTGGTGCCATGATGTAATCTTCTACACAGGTTTATTATCA 381
QY 361 CATGTTATGAATGGTGATTAAGAAAGCTTCATTTTAAATTTGGTGATGGCAGGTATGGGA 420
Db 382 CATGTTATGAATGGTGATTAAGAAAGCTTCATTTTAAATTTGGTGATGGCAGGTATGGGA 441

QY	421	TTGCCTGAAAGTGATTAATTCGCTTTTGGCGATGTGAATATAGACGTACATTTTAAAGT	480
Db	442	TTGCCTGAAAGTGATTAATTCGCTTTTGGCGATGTGAATATAGACGTACATTTTAAAGT	501
QY	481	TATAAACCTGATTAACGCAATTAATGACAAATATTAATTCGATCATCTGATTTTAA	540
Db	502	TATAAACCTGATTAACGCAATTAATGACAAATATTAATTCGATCATCTGATTTTCAA	561
QY	541	GATATTAAATGATTTTTCGATCATTCGAAGAAATGCGACATTAATGTTTAAAGGTAAT	600
Db	562	GATATTAAATGATTTTTCGATCATTCGAAGAAATGCGACATTAATGTTTAAAGGTAAT	621
QY	601	ATTGCTTGGGGTGATGATGAACATCTACGTAATAATTAAGCAGATGTTCCAAATTAAT	660
Db	622	ATTGCTTGGGGTGATGATGAACATCTACGTAATAATTAAGCAGATGTTCCAAATTAAT	681
QY	661	TATGGATTAAGATTCGGATGACATTTATGCTCAAAATATTAATTAAGGTAAGGT	720
Db	682	TATGGATTAAGATTCGGATGACATTTATGCTCAAAATATTAATTAAGGTAAGGT	741
QY	721	ACTGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	780
Db	742	ACTGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	801
QY	781	GGTGACCATACAGTTTAAATGCAATTAAGTGTAAATTCGATTAATTTAGAGAAGCTA	840
Db	802	GGTGACCATACAGTTTAAATGCAATTAAGTGTAAATTCGATTAATTTAGAGAAGCTA	861
QY	841	GATGTTCAAAATTAAGAAGCATTAAGAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	900
Db	862	GATGTTCAAAATTAAGAAGCATTAAGAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	921
QY	901	GAACCTACATTCGAATCAAGTTATTTAGATGATTAATGACACCATCCCAAGAGAAAT	960
Db	922	GAACCTACATTCGAATCAAGTTATTTAGATGATTAATGACACCATCCCAAGAGAAAT	981
QY	961	AGTGCTACATTTGAAACAGCAGCAAGAAATATCCACATAAGAAAGTTGTTGCGAGTATT	1020
Db	982	AGTGCTACATTTGAAACAGCAGCAAGAAATATCCACATAAGAAAGTTGTTGCGAGTATT	1041
QY	1021	CAACCAACACATTTCTCTAGAACACAGCAATTTTAAATGAATTTGCGAAGTTTAAAGT	1080
Db	1042	CAACCAACACATTTCTCTAGAACACAGCAATTTTAAATGAATTTGCGAAGTTTAAAGT	1101
QY	1081	AAACGAGATCGTATTTCTTATGTTGAAATTTTGGATCAATTAGAGAAATCTCGCGCA	1140
Db	1102	AAACGAGATCGTATTTCTTATGTTGAAATTTTGGCTCAATTAGAGAAATCTCGCGCA	1161
QY	1141	TTAACGATACAGATTTTAAATGATAAAATTTGAAGTGCATCGTTAAATTAAGAGATTTCT	1200
Db	1162	TTAACGATACAGATTTTAAATGATAAAATTTGAAGTGCATCGTTAAATTAAGAGATTTCT	1221
QY	1201	ATTAATGATTTAGAACAAATTTGATAATCTGTTATTTTATTTATGGGTGCGAGTGAATTT	1260
Db	1222	ATTAATGATTTAGAACAAATTTGATAATCTGTTATTTTATTTATGGGTGCGAGTGAATTT	1281
QY	1261	CAAAAATTAAGATGATTTATGATAAATTAAGCATGAAATTCGGTTTAA 1314	
Db	1282	CAAAAATTAAGATGATTTATGATAAATTAAGCATGAAATTCGGTTTAA 1335	
RESULT 9			
AAV74703			
ID	AAV74703 standard; DNA; 2424 BP.		
XX			
AC	AAV74703;		
XX			
DT	16-MAR-1999 (first entry)		
XX			
DE	Staphylococcus aureus contig SEQ ID #392.		
XX			
KN	Computer readable medium; vaccine; S.aureus infection; immunodetection;		

KN	cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KN	skin infection; surgical wound infection; scalded skin syndrome;
KN	toxic shock syndrome; ds.
OS	Staphylococcus aureus.
XX	Key Location/Qualifiers
PH	misc_feature 1141..1200
FT	/note="these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
XX	EP786519-A2.
PN	30-JUL-1997.
XX	07-JAN-1997; 97EP-00100117.
XX	05-JAN-1996; 96US-0009861P.
PR	(HUMA-) HUMAN GENOME SCI INC.
PA	Kunsch CA, Choi GH, Barash SC, Dillon PJ, Pannon MR, Rosen CA;
XX	WPI; 1997-374922/35.
DR	Polynucleotide(s) and proteins derived from Staphylococcus aureus - stored on computer readable medium and used in the production of anti-S.aureus vaccines.
PT	Claim 1; Page 1287-1288; 3271pp; English.
XX	This sequence represents one of 5191 Staphylococcus aureus DNA sequences of the invention. The DNA sequences are recorded on a computer readable medium, preferably selected from a floppy or hard disk, random access memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using the S.aureus DNA sequences allows putative functions to be assigned so that protein-encoding or regulatory regions of commercial, therapeutic or industrial importance can be obtained. Specifically, sequences which are likely to encode antigens have been identified and these polypeptides can be used in a vaccine composition against S.aureus infection. The polypeptides can also be used in a kit for the immunodetection of S.aureus in a sample. S.aureus is implicated in numerous human diseases, including cellulitis, eyelid infections, food poisoning, osteomyelitis, skin and surgical wound infections, scalded skin syndrome, toxic shock syndrome, etc. Organisms transformed with the DNA sequences can be used for recombinant production of the polypeptides. The new DNA sequences (and their fragments) are useful as primers or probes for isolating homologues of any of the S.aureus DNA sequences contained on the computer readable medium
SQ	Sequence 2424 BP; 839 A; 302 C; 462 G; 759 T; 0 U; 62 Other;
Query Match 95.1%; Score 1254; DB 2; Length 2424;	
Best Local Similarity 95.4%; Pred. No. 1.1e-236;	
Matches 1254; Conservative 0; Mismatches 60; Indels 0; Gaps 0;	
QY	1 ATGACACATCATCTTTTGTTCGGAATTAAGGTTCTGGCATGAGTTTCATTAGCACAATC 60
Db	627 ATGACACATCATCTTTTGTTCGGAATTAAGGTTCTGGCATGAGTTTCATTAGCACAATC 686
QY	61 ATGCGATGTTTAGGCATGAAGTTCAAGGATCGATATTCAGAACTACGATTTTACAGAA 120
Db	687 ATGCGATGTTTAGGCATGAAGTTCAAGGATCGATATTCAGAACTACGATTTTACAGAA 746
QY	121 GTTGCTCTTAGAAATAAGGGGATAAAAATATTTACCATTTGATGCTAATAACATAAGAA 180
Db	747 GTTGCTCTTAGAAATAAGGGGATAAAAATATTTACCATTTGATGCTAATAACATAAGAA 806
QY	181 GATATCGTATGTTATCAAGGTAATGATTCGCGAGTAGCCATGAGAAATAGTACGTGCA 240

Db 807 GATATGGTAGTTATACAAAGGTAATGCATTCGCGAGTAGCCATGAAGAAATAGTACGTGCA 866
Qy 241 CATCAATGAATAGATGTTGTAAGTTATAATGATTTTTTTAGGACAGATTAATGATCAA 300
Db 867 CATCAATGAATAGATGTTGTAAGTTATAATGATTTTTTTAGGACAGATTAATGATCAA 926
Qy 301 TATACCTCAGTAGCTGAACCTGGTCGACATGCTGAACCTTTTACACAGGTTTATATATCA 360
Db 927 TATACCTCAGTAGCTGAACCTGGTCGACATGCTGAACCTTTTACACAGGTTTATATCA 986
Qy 361 CATGTTATGAATGGTGATAAAGAACTTCATTTTTTAATTTGGTGATGGCACAGTATGGGA 420
Db 987 CATGTTATGAATGGTGATAAAGAACTTCATTTTTTAATTTGGTGATGGCACAGTATGGGA 1046
Qy 421 TTGCTGAAAGTGATTAATTCGCTTTTGGGCAATGTAATATAGACGTCACTTTTAAAGT 480
Db 1047 TTGCTGAAAGTGATTAATTCGCTTTTGGGCAATGTAATATAGACGTCACTTTTAAAGT 1106
Qy 481 TATAAACCCTGATTACGCAATATGACAAATATTTGATTTTCGATCATCCTGATTTATTTAA 540
Db 1107 TATAAACCCTGATTACGCAATATGACAAATATTTGNNNNNNNNNNNNNNNNNNNNNN 1166
Qy 541 GATATTATGATGTTTTTGTGATTCCTCAAGAAATGGCACATAATGTTTAAAAAGGTATT 600
Db 1167 NNN 1226
Qy 601 ATTGCTGGGGTGATGATGAACATCTACGTAAATTTGAAGCAGATGTTCCAAATTTATAT 660
Db 1227 ATTGCTGGGGTGATGATGAACATCTACGTAAATTTGAAGCAGATGTTCCAAATTTATAT 1286
Qy 661 TATGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTTCAAATTTACGGATAAGGT 720
Db 1287 TATGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTTCAAATTTACGGATAAGGT 1346
Qy 721 ACTGCTTTGATGCTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 780
Db 1347 ACTGCTTTGATGCTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1406
Qy 781 GGTGACCATACATGTTTTAAATGTCATAGCTGTAATTCGATTAATTTAGAGAGACTA 840
Db 1407 GGTGACCATACATGTTTTAAATGTCATAGCTGTAATTCGATTAATTTAGAGAGACTA 1466
Qy 841 GATGTTACAAATATTTAAAGAAAGCATTTAGAAACGTTTGGTGGTGGTGGTGGTGGTGGTGG 900
Db 1467 GATGTTACAAATATTTAAAGAAAGCATTTAGAAACGTTTGGTGGTGGTGGTGGTGGTGGTGG 1526
Qy 901 GAAACTCAATTCGAAATCAAGTTATTTGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 1527 GAAACTCAATTCGAAATCAAGTTATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1586
Qy 961 AGTCTCAATTTGAAACAGCAGCAAGAAATATCCACATAAAGAAAGTTGTTGCAGTATTT 1020
Db 1587 AGTCTCAATTTGAAACAGCAGCAAGAAATATCCACATAAAGAAAGTTGTTGCAGTATTT 1646
Qy 1021 CAACCAACACTTTCTCTAGAACACAGGCAATTTTAAATGATTTTGCAGAAAGTTTAAAGT 1080
Db 1647 CAACCAACACTTTCTCTAGAACACAGGCAATTTTAAATGATTTTGCAGAAAGTTTAAAGT 1706
Qy 1081 AAACGATGCTGTATCTTATGTAATGTAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1140
Db 1707 AAACGATGCTGTATCTTATGTAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1766
Qy 1141 TTAACGATACAGATTTTAAATGATAAATTCGAAGTGCATCGTTAATTAATGAAGTTCT 1200
Db 1767 TTAACGATACAGATTTTAAATGATAAATTCGAAGTGCATCGTTAATTAATGAAGTTCT 1826
Qy 1201 ATTAATGTTATGAAACAATTTGATAATGCTGTATTTTATTTATGGGTGCAGGTGATATT 1260
Db 1827 ATTAATGTTATGAAACAATTTGATAATGCTGTATTTTATTTATGGGTGCAGGTGATATT 1886
Qy 1261 CAAAAATTTACAAATTTAGATAAATTTAGGATCAATTTAGGATCAATTTAGGATCAATTTAGGATCA 1314
Db 1887 CAAAAATTTACAAATTTAGATAAATTTAGGATCAATTTAGGATCAATTTAGGATCAATTTAGGATCA 1940

RESULT 10

ABN92833
ID ABN92833 standard; DNA; 1329 BP.

XX
AC ABN92833;

XX
DT 24-JUL-2002 (first entry)

XX
Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:2296.

XX
Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
antibacterial; gene therapy; gene; ds.

XX
Staphylococcus epidermidis.

XX
US6380370-B1.

XX
30-APR-2002.

XX
13-AUG-1998; 98US-00134001.

XX
14-AUG-1997; 97US-0055779P.

XX
08-NOV-1997; 97US-0064964P.

XX
(GENO-) GENOME THERAPEUTICS CORP.

XX
Doucette-Stamm LA, Bush D;

XX
WPI; 2002-381255/41.

XX
P-PSDB; ABP40288.

XX
Novel isolated nucleic acid encoding a Staphylococcus epidermis

XX
polypeptide, useful for diagnosing and treating bacterial infections.

XX
Disclosure; SEQ ID NO 2296; 267pp; English.

XX
ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
frame (ORF) nucleic acid sequences which encode the amino acid sequences
given in ABP5124 to ABP37960. The S. epidermidis sequences have
antibacterial activity and can be used in gene therapy. The sequences can
also be used in the diagnosis and treatment of bacterial infections,
particularly S. epidermidis infections. The sequences can be used to
screen for compounds able to interfere with the S. epidermidis life cycle
or inhibit S. epidermidis infection. N.B. The sequence data for this
patent did not form part of the printed specification, but was obtained
in electronic format directly from the USPTO web site

XX
Sequence 1329 BP; 482 A; 180 C; 247 G; 420 T; 0 U; 0 Other;

Query Match 70.0%; Score 922; DB 6; Length 1329;

Best Local Similarity 81.4%; Pred. No. 1.3e-171;

Matches 1069; Conservative 0; Mismatches 245; Indels 0; Gaps 0;

Qy 1 ATGACACACTATCATTTTTCGGAATTAAGGTTCTGCGATGAGTTCATTAGCAACAATC 60

Db 16 ATGACACACTATCATTTTTCGGAATTAAGGTTCTGCGATGAGTTCATTAGCAACAATC 75

Qy 61 ATGCATGATTTAGGACATGAAGTTCAAGGATTCGGATATTCGAGAACTACGTTATTACAGAA 120

Db 76 ATGCATGACCTCGTTCATGAAGTTACAAGGTTTCAGACATAGATCATACGTTTATACAGAA 135

Qy 121 GTTGTCTTTAGAAATTAAGGGGATAAAAATTAATTCATTTGATGCTAATAACATAAAGAA 180

Db 136 GTTGTCAATTAAGAAATTAAGGGATTAATAATTTTACCTTTTCATGCAATAATATTACAAA 195

Qy 181 GATATGGTAGTTATACAGGTAATGATTCGCCAGTAGCCATGAGAAATAGTACGTGCA 240

Db 196 GAAATGGTTGTCATCCCAAGGTAATGCAATTCCTTCGTAATCATGAAGAAATTTGTAGGGCA 255

Qy 241 CATCAATTCGAAATTTAGATGTTGTAAGTTATAATGATTTTTTTAGGACAGATTTATTGATCAA 300

Db 256 CATGAATTAAGCTGTGATATTTATATAAATATCATGACCTTTCTTGCTCATGTTATTAATCAA 315
QY 301 TATACCTTCAGTAGCTGTAACCTGTCACATGGTAAACCTTACAAACAGGTTTATATATCA 360
Db 316 TATACATCTGCTGCTTACAGGTGACACGCTAAACGCTCAACACTGGTTTGTGTCA 375
QY 361 CATGTTATGAATGGTGTATATAAAGACTTCATTTTAAATTTGTTGATGGACAGGTATGGGA 420
Db 376 CATGTGATGAATGGATATAAAGACATCATTCCTAATCGGTGATGGTACAGGTATGGGA 435
QY 421 TTGCTCAAGAGTGAATTTTTCGCTTTTGGAGCAGTGAATATAGACGTCACCTTTTAAAGT 480
Db 436 CTACAGGAAGTGAATTTTTCGCTTTTGGAGCTTGTGAATACCGTAGGCATCTTCTTAAGT 495
QY 481 TATAAACCCTGATACGCAATTTATGACAAATATTTGATTTTCGATCATCTCTGATTTTAA 540
Db 496 TATCATCCGATTTATGCAATTTATGACTAATATTTGATTTTGACCATCCAGATTTTAA 555
QY 541 GATATTAATGATGTTTTTGTGATTCACAGAAATGGCACAATAATGTTTAAAGGTTAT 600
Db 556 AATATCGATGATGTTTATGATGATTCACATATGCGCTTAAATGTTTAAAGAGGAAT 615
QY 601 ATTGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 616 ATTGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 675
QY 661 TATGGAATTTAAGATTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 676 TATGGCTTTAAGAAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 735
QY 721 ACTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 736 ACGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 795
QY 781 GGTGACCATACGCTTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 796 GGAATCACACACATTCCTAAACGCACTTCCTGATATGCAATTTAGTTATTTAGAAAACATG 855
QY 841 GATGTTACAAATTTAAGAAAGCATTTAGAAAGCTTTGCTGCTGTTTAAACGCTGTTCAAT 900
Db 856 AATGTAGAAATTTAAGAAAGCATTTGATGATGATGATGATGATGATGATGATGATGATGAT 915
QY 901 GAACTACCAATTCGAATCAAGTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 916 GAAACAAAAGTATCAATCAAGTAAATAGTAGACGATTTATGACACCATCCAAAGAGAAAT 975
QY 961 AGTGCTACCAATTTGAAACAGACGACGAAAGAAATATCCACATAAAGATTTGTCAGTATTT 1020
Db 976 AGTGCTACCAATTTGAGACGACGCGGAAAGAAATATCCGCAAAAGATTTGTTGCGGTTTT 1035
QY 1021 CAACCAACACTTCTCTAGAACACAGCGCATTTTAAATGAATTTGCAGAAAGTTTAAAGT 1080
Db 1036 CAACCAACACTTCTCTAGAACACTCAAGCATTTTAAACGAATTTGCTGAGAGTTTAAAGT 1095
QY 1081 AAAGCAGATCGTGTATTTCTTATGTAATTTTGGATCAATTAGAGAAATATCTGGCGCA 1140
Db 1096 AAGGCAGACCAAGTATTTTATGTAATTTTGGTCAATTAAGAGAAATATCGGAGAT 1155
QY 1141 TTAACGATACAGATTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Db 1156 TTAACATAGAGATTTTAAATCAATTCGTTATGACGCTCGACGCTTATAGATGAAATAGT 1215
QY 1201 ATTAATGATTTAGAACATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db 1216 AATGATGATTTAGAAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1275
QY 1261 CAAAAATTTACAAATGCATATTTAGATAAATTTAGGCATGAAAAATGCGTTTAA 1314
Db 1276 CAAAAGTTACTAAAGCTTATTCGAAAATTTAGCGTAAATGATTTTAA 1329

RESULT 11
ACA47124

ID ACA47124 standard; DNA; 1311 BP.
XX ACA47124;
XX 19-JUN-2003 (first entry)
XX Prokaryotic essential gene #28781.
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX Staphylococcus epidermidis.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR P-PSDB; ABU43254.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX Claim 14; SEQ ID NO 34994; 1766pp; English.
PS The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 1311 BP; 473 A; 181 C; 243 G; 414 T; 0 U; 0 Other;

Query Match		69.5%;	Score 915.8;	DB 7;	Length 1311;
Best Local Similarity		81.2%;	Pred. No. 2.2e-170;		
Matches 1064;		Conservative	0;	Mismatches 247;	Indels 0; Gaps 0;
QY	1	ATGACACATCATCTTTTGTGCGGAATTAAGGTTCTGCGCATGAGTTCATTAGCACAATC	60		
DB	1	ATGACACATCATCTTTTGTGCGGAATTAAGGTTCTGCGCATGAGTTCATTAGCACAATC	60		
QY	61	ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTTATTTACAGAA	120		
DB	61	ATGCATGACCTCGTCATGAGTCAAGGTTCAAGGATCGGATATTGAGAACTACGTTATTTACAGAA	120		
QY	121	GTTCCTCTTAGAATAAGGAGTAAATAATATACCAATTTGATGCTTAATAACATAAAGAA	180		
DB	121	GTTCATTAAGAAATTAAGGATTAATAATTTTACCTTTTGTGCAATAATATTAACAAA	180		
QY	181	GATATGTAGTTATACAGGTAATGCAATTCGCGAGTACGCAATGAGAAATAGTAGTGCA	240		
DB	181	GAAATGGTTGTCATCAAGGTAATGCAATTCCTGATAATCATGAAGAAATTTGTAGGGCA	240		
QY	241	CATCAATTTGAATTTAGATGTTTGAAGTTATATGATTTTTTTAGGACAGATTTTGCATCA	300		
DB	241	CATGAATTAAGCTTGATTTATTAATAATATCATGATTTCTGGTCATGTTATATAATCA	300		
QY	301	TATACCTCAGTGTGTAACCTGGTGCAATGGTAAACCTTTTACACAGGTTTATATCA	360		
DB	301	TATACATCTGTTGCTGTACAGGTGACACGGTAAACCTTCAACCTGGTTGTGTCA	360		
QY	361	CATGTTATGATGTTAATAAAGACCTTCATTTTTTAATTTGGTGATGGCACAGGTATGGA	420		
DB	361	CATGTGATGAATGAGATTAATAAAGACATCTTCCTAATCGGTGATGGTACAGGTATGGA	420		
QY	421	TTGCTGAAAGTGATTTATTCGCTTTTCAAGGATGTAATATAGACGTCACCTTTTAACT	480		
DB	421	CTACAGGAGTACTATTTTCTGCTTTTCAAGGCTTGTAATACCGTAGGCATCTTTAAGT	480		
QY	481	TATAAATCTGATTAAGCAATTAATGACAAATATGATTTTTCGATCATCTGATTTTAA	540		
DB	481	TATCATCCAGATTTATGCAATTAATGATTAATTTGATTTTGACCATCCAGATTTTAA	540		
QY	541	GATATTAATGATTTTTTGTGATTCACAGAAATGGCACATAATTTTAAAGAAAGTAT	600		
DB	541	AATATCGATGATTTTATGATGATTCACCAATATGATGCAATTAATTTTAAAGAAAGTAT	600		
QY	601	ATTGCTTTGGGTGATGATGACATCTAGTAAATTTGAAGCAGATGTTCCAAATTTATAT	660		
DB	601	ATTGCTTTGGGTGATGATGATTAATTTAGTAAATTTAGATGTTGATTTCCGGTTTATAT	660		
QY	661	TATGATTTAAGATTTCCGATGACATTTTATGCTCAAAATATTTCAAAATACGGATAAAGT	720		
DB	661	TATGCTTTAAGAAACAGATGACATCTATGCTAAATAATTTCAAAATTTACTGAAAGGT	720		
QY	721	ACTGCTTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	780		
DB	721	ACGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	780		
QY	781	GGTGACACATAGTTTAAATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT	840		
DB	781	GGAAATCACAACTTTTAAAGCACTTCTGCTAATAGCAATTTAGTATTTTAGAAGACATG	840		
QY	841	GATGTTACAATATTAAGAGGATTTAGAAAGTTTGGTGGTAAAGCTGCTTTCAAT	900		
DB	841	AATGTAGAAATATTAAGAGGATTTAGAAAGTTTGGTGGTAAAGCTGCTTTCAAT	900		
QY	901	GAAGCTCAATTTGAAATCAAGTTATTTAGATGATTTATGACACATCCAGAGAAAT	960		
DB	901	GAACAAAAGTATCAATCAAGTATTTAGTACGATTTATGACACATCCAGAGAAAT	960		
QY	961	AGTGCTCAATTTGAACAGCAGCAAGAAATATCCACATAAAGAGTTGTTGCAGTATTT	1020		
DB	961	AGTGCTCAATTTGAACAGCAGCAGCAAGAAATATCCGCAAAAGAGTTGTTGCAGTATTT	1020		
QY	1021	CAACACACACTTTCTCTAGAACACAGGCAATTTTAAATGATTTTGCAGAAAGTTAAGT	1080		

Db	1021	CAACCATACGTTCTCAAGAACTCAAGCAATTTTAAACGAATTTGCTGAGAGTTTAAGT	1080	
Qy	1081	AAAGCAGATCGTGTATTTCTTATCTGAAATTTTTCGATCAATTTAGAGAAAATTA	1140	
Db	1081	AAGCGACCAAGTATTTTATGTAATATTCGGTTCAATAAGAGAAAATACGGAGAT	1140	
Qy	1141	TTAACGATACAAGATTTAAATTTGATAAAATTTGAAGGTGCATCGTTAATTAATGAAGATCT	1200	
Db	1141	TTAACTATAGAAGATTTAAATTCGATTCGATTCGACGGCTTATAGATGAAAATAGT	1200	
Qy	1201	ATTAATGTTATTAAGCAATTTGATAATGCTGTTATTTTATTTATGCGTGCAGGTGATATT	1260	
Db	1201	ATTGATGTTATGAGAAAATTTGATAACGCTGTGTAATTTTATTTATGGCGCTGGTGACATT	1260	
Qy	1261	CAAAAATTTACAAAATTCATATTTTAGATAAATTTAGGCATGAAAATGCGTTT	1311	
Db	1261	CAAAAGTTACTAAAAGCTTATTCGAAAAATTTAGCGCTAAAANAATGATTTT	1311	
RESULT 12				
AAV80065				
ID	AAV80065 standard; DNA; 660 BP.			
XX	AAV80065;			
XX	17-MAR-1999 (first entry)			
XX	Partial nucleotide sequence of the MurC gene.			
XX	MurC gene; UDP-N-acetylmuramate:L-alanine ligase; MurC polypeptide;			
KW	bacterial; infection; H. pylori; cancer; ulcer; gastritis; vaccine;			
KW	immunogen; drug; genetic immunisation; ds.			
OS	Staphylococcus aureus.			
FH	Key	Location/Qualifiers		
CDS	2..660	/tag= a product= "partial MurC polypeptide"		
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				
FT				


```
QY 610 GGTGATGATGAACATCTACGTAAATTTGAAGCAGATGTTCCAAATTTATTATTATGCAATTT 669
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 658 GGTGATGATGAATATCTTCGCCAGTTAGAAATCAGAAGTGCCAGTTATTATTATGGCGTC 717
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 670 AAAGATTCGGATGACATTTATGCTCAAAATATTCAAAATTTACGGATAAAGGTACTGCTTTT 729
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 718 AGCGAAGAGGATGATATCCAGCCCCGAAATATTCAACGAACCAACCGAAGGCTCATCTTTT 777
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 730 GATGTTATGTTGGATGAGTGTGAGTTTATGATCATCTTCTGCTCCACAATATGTTGACCAT 789
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 778 GATGTTTATCAAGAGCTGATTTTGTAGTCAATTTGCTTACCGCATTTGGCCATCAC 837
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 790 ACAGTTTTAAATGCAATAGCTGTAATGCGATTAGTTATTTAGAGAAGCTAGATGTTACA 849
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 838 AATATCATGAATGCGCTAGGTGTGATGCTGTGGCTTATTTGAAAAAACTTGATATGCAA 897
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 850 AATATTAAGAGCAATTAAGAACGTTTGGTGGTGTAAACGTCGTTTCAATGAAGACTACA 909
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 898 AAAGTCGCAAGAGAAATGCTAAGTTTAAAGGTGTGAACGTCGTTTACGCGAGAAAAA 957
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 910 ATTGCAATCAAGTTATTGTAGATGATTATGCACACCATCCAAAGAGAAATTAGTGTCTACA 969
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 958 GTCAGTGACATGATTTATTTGTAGTATTTATGCGCACCATCCAGCTGAATTTAAAGCAAG 1017
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 970 ATTGAACACAGCACGAAAGAAATATCCACATAAAGAGTTGTTGCAGTATTTCAACCACAC 1029
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1018 ATTGATGGCGCCGCCAAAATATCCTGCACAAAGAAATTTATGCTGTCTCCAGCCACAT 1077
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1030 ACTTCTCTAGAACACAGGCAATTTTAAATGAATTTGCAGAAAGTTTAAAGTAAAGCAGAT 1089
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1078 ACATTTACCGAACAAATGTCCTTAATGGATGAATTTGCTGAAGCACTGGATTTGGCAGAT 1137
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1090 CGTGTATCTTATGTGAATTTTGGATCAATTAGAGAAATACTGCGGCATTAACGATA 1149
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1138 GAAGTATCTTATGTAATTTTGGCTCTGCGGTGAACACAAAGGCGAGGTACGCAAT 1197
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1150 CAAGATTTAATGTATAAATTT---GAAGGTGCATCGTTAAATTAATGAAGATTTCTATTAT 1206
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1198 GAAGATTTAGTGAAAAAATTCAAAAGGTGGCAAGTAAATTCGGAAGATAATGTGTGCG 1257
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1207 GTATTGAACAAATTTGATATGCTGTATTTTATTTATGGGTGCAAGGTGATATTCAAAAA 1266
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1258 CCTTACTAGATTTTGAATAATGCAGAGGTGCTTTTATGGGTGCTGGCGAGCTTCAGAAA 1317
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1267 TTACAAATGCATATTAGATATAATTAGGCCATGAARAT 1305
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1318 TTGAAACAGCCTACGAAACAAATTAAGTAACAACT 1356
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: June 6, 2004, 00:10:04
Job time : 642 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2004, 23:46:19 ; Search time 134 Seconds
(without alignments)
5458.398 Million cell updates/sec

Title: US-10-712-713-1

Perfect score: 1318

Sequence: 1 atgacacactatcatttgt.....gaaaatggttttaagctt 1318

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A-COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B-COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A-COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B-COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PTUS-COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1314	99.7	2424	4	US-08-956-171E-392
2	922	70.0	1329	4	US-09-134-001C-2296
3	570.4	43.3	619	4	US-08-936-165A-179
4	400.8	30.4	11864	4	US-08-961-527-61
5	399.2	30.3	2715	4	US-08-940-572-1
6	389.6	29.6	888	3	US-08-714-918-61
7	389.6	29.6	888	3	US-09-265-315-61
8	389.6	29.6	888	3	US-09-265-315-61
9	389.6	29.6	888	3	US-09-266-417-61
10	389.6	29.6	888	4	US-09-528-709-61
11	389.6	29.6	888	4	US-09-527-745-61
12	370.2	28.1	1267	3	US-08-961-083-115
13	370.2	28.1	1267	4	US-09-536-784-115
14	354.4	26.9	873	4	US-09-107-532A-3260
15	310.4	23.6	1825	4	US-08-940-572-3
16	138.6	10.5	1479	4	US-09-328-352-897
17	105.4	8.0	640681	4	US-09-790-988-1
18	84.4	6.4	1467	4	US-09-543-681A-2779
19	78.4	5.9	1830121	4	US-09-557-884-1
20	78.4	5.9	1830121	4	US-09-643-990A-1
21	76.2	5.8	1230025	4	US-09-138-452A-1
22	72.8	5.5	5361	3	US-08-973-462-2
23	72.8	5.5	6152	3	US-08-973-462-1
24	58.8	4.5	1485	4	US-09-543-681A-601
25	55.2	4.2	8920	2	US-08-446-855A-1
26	55.2	4.2	8920	3	US-09-150-741-1
27	54.4	4.1	1830121	4	US-09-557-884-1

C 28	54.4	4.1	1830121	4	US-09-643-990A-1	Sequence 1, Appli
C 29	54	4.1	15016	4	US-09-601-198-60	Sequence 60, Appli
C 30	53.8	4.1	2223	1	US-08-257-073-4	Sequence 4, Appli
C 31	53.6	4.1	1891	3	US-08-973-462-3	Sequence 3, Appli
C 32	53.4	4.1	1482	4	US-09-540-236-792	Sequence 792, App
C 33	53.4	4.1	19619	4	US-09-596-002-14	Sequence 14, Appli
C 34	52.8	4.0	5340	4	US-09-627-122-21	Sequence 21, Appli
C 35	52.4	4.0	14066	4	US-09-601-198-56	Sequence 56, Appli
C 36	50.8	3.9	1374	4	US-09-134-001C-1447	Sequence 1447, Ap
C 37	50.2	3.8	1416	4	US-09-543-681A-1533	Sequence 1533, Ap
C 38	50	3.8	9636	1	US-08-323-170B-1	Sequence 1, Appli
C 39	50	3.8	9636	4	US-08-954-441-1	Sequence 1, Appli
C 40	49.8	3.8	2606	2	US-08-568-459A-7	Sequence 7, Appli
C 41	49.8	3.8	2606	2	US-08-487-826B-7	Sequence 7, Appli
C 42	49.8	3.8	2606	4	US-09-210-288-7	Sequence 7, Appli
C 43	49.6	3.8	520	4	US-08-961-527-344	Sequence 344, App
C 44	49.4	3.7	1692	4	US-09-601-198-63	Sequence 63, Appli
C 45	49	3.7	5652	4	US-09-601-198-75	Sequence 75, Appli

ALIGNMENTS

RESULT 1

US-08-956-171E-392
; Sequence 392, Application US/08956171E
; Patent NO. 6593114

GENERAL INFORMATION:

APPLICANT: Charles Kunsch
Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,171E

FILING DATE: 20-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861

FILING DATE: January 5, 1996

APPLICATION NUMBER: 08/781,986

FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB248P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (240) 314-1224

TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 392:

SEQUENCE CHARACTERISTICS:

LENGTH: 2424 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 392:

US-08-956-171E-392

Query Match		99.7%	Score 1314;	DB 4;	Length 2424;
Best local Similarity		100.0%;	Pred. No. 6.4e-296;		
Matches 1314;		Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	ATGACACACTATCATTTTGTTCGGAAATTAAGAGTTCTGGCATGAGTTCAATTAGCACAAATTC	60		
Db	627	ATGACACACTATCATTTTGTTCGGAAATTAAGAGTTCTGGCATGAGTTCAATTAGCACAAATTC	686		
QY	61	ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTGACTGATTTTACAGAA	120		
Db	687	ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTGACTGATTTTACAGAA	746		
QY	121	GTTCCTCTTAGAAATAAGGGATAAAAAATTACCAATTTGATCTAAATAACATAAAGAA	180		
Db	747	GTTCCTCTTAGAAATAAGGGATAAAAAATTACCAATTTGATCTAAATAACATAAAGAA	806		
QY	181	GATATCGTAGTTATACAAGGTAAATGCAATTCGGAGTAGGCCATGAAGAATAAGTAGTCGCA	240		
Db	807	GATATCGTAGTTATACAAGGTAAATGCAATTCGGAGTAGGCCATGAAGAATAAGTAGTCGCA	866		
QY	241	CATCAATTCGAAATTTAGATGCTGTAAAGTTATATGATTTTATAGGACAGATTTATGATCAA	300		
Db	867	CATCAATTCGAAATTTAGATGCTGTAAAGTTATATGATTTTATAGGACAGATTTATGATCAA	926		
QY	301	TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTCAACAAGGTTTATATATCA	360		
Db	927	TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTCAACAAGGTTTATATATCA	986		
QY	361	CATGTTATCAATGGTGATATAAAGACCTTCATTTTTTAAITGGTGACAGATTTATGATCAA	420		
Db	987	CATGTTATCAATGGTGATATAAAGACCTTCATTTTTTAAITGGTGACAGATTTATGATCAA	1046		
QY	421	TTGCCTGAAAGTGATTTATTCGCTTTTTCAGGCATGTGAATATAGACGCTCACTTTTAAAGT	480		
Db	1047	TTGCCTGAAAGTGATTTATTCGCTTTTTCAGGCATGTGAATATAGACGCTCACTTTTAAAGT	1106		
QY	481	TATAAACTCGATTACGCAATTTATGACAAATATTGATTTTCGATCATCTCGATTTATTTAAA	540		
Db	1107	TATAAACTCGATTACGCAATTTATGACAAATATTGATTTTCGATCATCTCGATTTATTTAAA	1166		
QY	541	GATATTTAATGATGTTTTTGATGCAATCCAGAAATGGCACATAATGTTTAAAAAGGTAAT	600		
Db	1167	GATATTTAATGATGTTTTTGATGCAATCCAGAAATGGCACATAATGTTTAAAAAGGTAAT	1226		
QY	601	ATTGCTTGGGGTGATGATGAACATCTACGTAAAAATGACAGATGTTCCAAATTTATAT	660		
Db	1227	ATTGCTTGGGGTGATGATGAACATCTACGTAAAAATGACAGATGTTCCAAATTTATAT	1286		
QY	661	TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTCGGATAAAGGT	720		
Db	1287	TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTCGGATAAAGGT	1346		
QY	721	ACTGCTTTTGATGTATGTGGATGGTGAGTTTATGATCAGTTCTCTGCTCCACAATAT	780		
Db	1347	ACTGCTTTTGATGTATGTGGATGGTGAGTTTATGATCAGTTCTCTGCTCCACAATAT	1406		
QY	781	GGTGACCATACAGTTTTTAAATGCAATTAGCTGTAATTTGCGATAGTTTATTTAGAGAGCTA	840		
Db	1407	GGTGACCATACAGTTTTTAAATGCAATTAGCTGTAATTTGCGATAGTTTATTTAGAGAGCTA	1466		
QY	841	GATGTTTCAAAATTTAAAGAAGCAATTAGAAACGTTTTGGTGGTTTAAACGCTGTTTCAAT	900		
Db	1467	GATGTTTCAAAATTTAAAGAAGCAATTAGAAACGTTTTGGTGGTTTAAACGCTGTTTCAAT	1526		
QY	901	GAAACTACAAATTCGAATCAAGTTATTGTAGATGATTATGCAACCATCCAGAGAAATTT	960		
Db	1527	GAAACTACAAATTCGAATCAAGTTATTGTAGATGATTATGCAACCATCCAGAGAAATTT	1586		
QY	961	AGTGCTACAAATTCGAAACAGCAGCAAGAAATATCCACATAAAGAGAGTTTGTGCGATTTT	1020		
Db	1587	AGTGCTACAAATTCGAAACAGCAGCAAGAAATATCCACATAAAGAGAGTTTGTGCGATTTT	1646		
QY	1021	CAACCCACACTTTCTCTGAAACACAGGCAATTTTTTAAATGAAATTTGCGAAGGTTTAAAGT	1080		

```

1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
251
```

Query Match	70.0%;	Score 922;	DB 4;	Length 1329;
Best Local Similarity	81.4%;	Pred. No. 5.1e-205;		
Matches 1069;	Conservative	0;	Mismatches 245;	Indels 0;
				Gaps 0;
QY	1	ATGACACATCATATTTGTTCGGAAATTAARAGGTTCTCGGCATGAGTTTCATTAGCACAAATC	60	
Db	16	ATGACACATCATATTTGTTCGGAAATTAARAGGTTCTCGGCATGAGTTTCATTAGCACAAATC	75	
QY	61	ATGCATGATTTAGGACATGAAGTTTCAAGGATCGGATATTCGAAACTACGTATTATTCAGAA	120	
Db	76	ATGCATGACCTCGTCACTGAAGTACAAAGGTTCCAGACATAGAATCATACGTTTTTCAGAA	135	
QY	121	GTTCGCTCTTAGAATAAAGGGGATAAAAATAATACCAATTTGATGCTTAATACATAAAGAA	180	
Db	136	GTTCGCTTAAGAATAAAGGGATTAATAATTTTACCTTTTGTATGCAAAATATATTACAAA	195	
QY	181	GATATGGTAGTTATACAAGGTAAATGCATTCGCGAGTAGCCATCAAGCAAAATAGTACGTGCA	240	
Db	196	GAATGGTTGTCAATCCAAAGTAAATGCATTTCTGTATAATCATGAAGAAATGTTAGGGCA	255	
QY	241	CATCAATTCGAATTAGATGCTTGTAATGATTTTTTTAGGCACAGATTAATGATCAA	300	
Db	256	CATGAATTAAGCGTTGATATTATAAAATATCATGACTTTCTTGGTCATGCTTAATAATCAA	315	
QY	301	TATACCTTCAGTAGCTGTACTCGTGACATGCTGTAAACCTTCTACCAACAGGTTTTATTATCA	360	
Db	316	TATACATCTGTTGCTGTACAGGTGCACACCGGTAAAACGCTCTCAACTGGTTTGTGTCA	375	
QY	361	CATGTTATCAATGGTGATAAAAAGACTTCATTTTTAAATTTGGTGATGGCACAGGTATGGGA	420	
Db	376	CATGTGATGAATGGGAGATAAAAAGACATCATTCCTTAATCGGTGTAGTGTACAGGTATGGGA	435	

```
QY 421 TTGCTGAAAGTGATTAATTTGGCTTTTGGAGCGATGTAATATAGACGCTCTTTTAAAGT 480
Db 436 CTACCGAGAGTGACATTAATTTGGCTTTTGGAGCTGTGTAATACCGTAGGCACTTCTTAAGT 495
QY 481 TATAAACCTGATTAACCAATATGACAAATATGATTTGATTCGATCATCTCGATTAATTTAAA 540
Db 496 TATCATCCAGATATGCAATATGATTAATGATTAATTTGACCATCCAGATTAATTTAAA 555
QY 541 GATATTAATGATGTTTTTGGATGATTTCCAGAAATGGCACAATATGTTTAAAAAGGTATT 600
Db 556 AATATCGATGATGTTTATGATGATTTCCACATATGCGGTTAAATGTTTAAAGAGGAATT 615
QY 601 ATTGCTTGGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 616 ATTGCTTGGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 675
QY 661 TATGATTTAAAGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 676 TATGCTTTAAAGAAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 735
QY 721 ACTGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 736 AGGCAATTGATGATATATTAAGGCGAATTTTATGATCAATCTTATCCCCACCAATAT 795
QY 781 GGTGACCATACAGTTTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 796 GGAATACCAACATTTCTAAACGCACTGCTGTAATAGCAATAGTATTTTAAAGAACATG 855
QY 841 GATGTTACAAATTTAAAGAACATTTAGAAAGCTTTGGTGGTGTGTTTAAAGCTGTTTCAAT 900
Db 856 AATGTAGAAAATTTTAAAGAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 915
QY 901 GAACTACAAATTGCAAAATCAAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 916 GAAACAAAGTATCAAAATCAAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 975
QY 961 AGTGCTACAATTTGAACAGCAGCAAGAAATATCCACATAAAGAGTGTGTCAGTATTT 1020
Db 976 AGTGCTACAATTTGAACAGCAGCAGGAAATATCCGAAAGAGTGTGTCAGTATTT 1035
QY 1021 CAACACACATTTCTCTAGAACACAGGCAATTTTAAATGAAATTTGCAAGAAAGTTTAAAGT 1080
Db 1036 CAACACACATGCTTCTCAAGAACTCAAGCATTTTAAAGCAATTTGCTGAGAGTTTAAAGT 1095
QY 1081 AAAGCAGATCGTATTTCTTATGTAATTTTGAATTTTGAATTTGATGATGATGATGATGATGAT 1140
Db 1096 AAGGAGACCAAGTATTTTATGTAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 1155
QY 1141 TTAACGATACAAAGATTTTAAATGATAAATTTGAAGGTGATCGTAAATTAATGAAGTTCT 1200
Db 1156 TTAACATAGAGATTTTAAATGATAAATTTGAAGGTGATCGTAAATTAATGAAGTTCT 1215
QY 1201 ATTAATGATTTAGAACAAATTTGATAATCTGTTATTTTATTTATGGGTGCGGTGATATT 1260
Db 1216 ATTGATGATTTAGAAAATTTGATAACGCTGTAATTTTATTTATGGCGCTGGTGACATT 1275
QY 1261 CAAAAATTCAAAATGCAATTTTATGATAAATTTAGGCATGAAAAATGCGGTTTAA 1314
Db 1276 CAAAAGTTACTAAAAGCTTATTTCCGAAAATTTAGGCGTAAAAAATGAATTTTAA 1329
```

RESULT 3

```
US-08-936-165A-179/C
; Sequence 179, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
```

```
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,165A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,032
; FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 179:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-936-165A-179
```

```
Query Match 43.3%; Score 570.4; DB 4; Length 619;
Best Local Similarity 98.0%; Pred. No. 1.6e-123;
Matches 588; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 716 AAGGTACTGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 775
Db 616 ACGAGACTGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 557
QY 776 AATATGCTGACCATACATGATTTTAAATGCTAGCTGTAATTTGCCATTAGTTATTATTAGAGA 835
Db 556 AATATGCTGACCATACATGATTTTAAATGCTAGCTGTAATTTGCCATTAGTTATTATTAGAGA 497
QY 836 AGCTAGATGTTACAAATATTTAAAGAAAGCATTTAGAAACGTTTGGTGGTGTGTTAAACGTCGTT 895
Db 496 AGCTAGATGTTACAAATATTTAAAGAAAGCATTTAGAAACGTTTGGTGGTGTGTTAAACGTCGTT 437
QY 896 TCAATGAAACTACAATTTGCAATCAAGTTATTGATGATGATGATGATGATGATGATGATGATGATGAT 955
Db 436 TCAATGAAACTACAATTTGCAATCAAGTTATTGATGATGATGATGATGATGATGATGATGATGATGAT 377
QY 956 AAATAGTCTCAATTTGAAACAGCAGCAAGAAATATCCACATAAAGAGTTGTTTCAG 1015
Db 376 AAATAGTCTCAATTTGAAACAGCAGCAAGAAATATCCACATAAAGAGTTGTTTCAG 317
QY 1016 TATTTCAACACACACTTTCTCTAGAAACAGAGCATTTTAAATGAATTTGCAAGAAAGTT 1075
Db 316 TATTTCAACACACACTTTCTCTAGAAACAGAGCATTTTAAATGAATTTGCAAGAAAGTT 257
QY 1076 TAAAGTAAAGCAGATCGTGTATTCTTATGTAATTTTGGATCAATTTAGAGAAAATACTG 1135
```

Db	256	TATGTAAGCAGATCGTGTATCTTATGTGAAATTTTGGCTCAATTAGAGAAAATTCGTG	197
Qy	1136	GCGCATTAACGATACAGATTTTAATTCATAAAATTTGAAGGTGCATCGTTAATTAATGAAG	1195
Db	196	GCGCATTAACGATACAGATTTTAATTCATAAAATTTGAAGGTGCATCGTTAATTAATGAAG	137
Qy	1196	ATTC-TATTAAATGATTAGAACAAATTTGATAATGCTGTTATTTTATTATGGGTGCAGGT	1254
Db	136	ATTCTTATTAAATGATTAGAACAAATTTGATAATGCTGTTGTTTATTATGGGTGCAGGT	77
Qy	1255	GATATTCAAAATTTACAAATTCATATTAGATAAAATTTAGCATGAAAAATGCGTTTAA	1314
Db	76	GATATTCAAAATTTACAAATTCATATTAGATAAAATTTAGCATGAAAAATGCGTTTAA	17
RESULT 4			
US-08-961-527-61/c			
; Sequence 61, Application US/08961527			
; Patent No. 6420135			
; GENERAL INFORMATION:			
; APPLICANT: Charles Kunsch			
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences			
; NUMBER OF SEQUENCES: 391			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Human Genome Sciences, Inc.			
; STREET: 9410 Key West Avenue			
; CITY: Rockville			
; STATE: Maryland			
; COUNTRY: USA			
; ZIP: 20850			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage			
; COMPUTER: HP Vectra 486/33			
; OPERATING SYSTEM: MSDOS version 6.2			
; SOFTWARE: ASCII Text			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/961,527			
; FILING DATE:			
; CLASSIFICATION: 424			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER:			
; FILING DATE:			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Brookes, A. Anders			
; REGISTRATION NUMBER: 36,373			
; REFERENCE/DOCKET NUMBER: PB340P1			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (301) 309-8504			
; TELEFAX: (301) 309-8512			
; INFORMATION FOR SEQ ID NO: 61:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 11864 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: linear			
; US-08-961-527-61			

	Query Match	30.4%;	Score 400.8;	DB 4;	Length 11864;
	Best Local Similarity	59.2%;	Prod. No. 9.3e-84;		
	Matches 760;	Conservative 0;	Mismatches 512;	Indels 12;	Gaps 4;
QY	10	TATCATTTTGTCCGAATTAAAGGTTCTGGCATCAGTTTCATTAGCACAAATCATGCATGAT	69		
Db	1749	TATCATTTTATCCGAATTAAAGGATCAGGATCAGTGCCCTTGGCTTGTGTCACACG	1690		
QY	70	TTAGGCATCAAGTTTCAAGGATCCGATATTGAGAACTACGTTATTTCAGAAAGTTGCTTT	129		
Db	1689	ATGGGGCACAAGTTTCAGGGATCAGATGTTTGAAGAAAGTACTACTTTATCCCAACCGGTCCT	1630		
QY	130	AGAAATTAAGGGGATAAAAAATTTACCATTTTGATGCTTAATAACATAAAGAGGATATGGTA	189		
Db	1629	GAGCAGCGAGGAATTAACCATTTCTCTTTTGATGAATAAAATCTCAGACGGTGATATGGAA	1570		

QY	190	GT	TATACAAAGTAAATGCAATTTTTCGCGAGTAGCCATGAAGAAATAAGTAGCGTGCACATCAA	246
Db	1569	ATT	ATCGCTCGAATAATGCGCTTCGTCACAGATAACAACGTCGAAATTCCTATGCGGACCAA	1510
QY	247	TTG	AAATTTAGATGTTGTAAGTTATAATGATTTTTTTAGGACAGATTTATGATCAATATACT	306
Db	1509	AA	TGGTATCAGCTACAAACGTTACATGAGTTTCTAGGTAGCTTTATGCGTGACATTTGTT	1450
QY	307	TC	AGTAGCTGTAACTGGTGCAATGTTTAACTTCTACAAACAGGTTTATTTATCACAATGTT	366
Db	1449	AG	CATGGGAGTAGCAGGAGCACATGGAAAACTTCAAACGACAGGTATGTTGTCTAAGTC	1390
QY	367	AT	GAATCGGTGATAAAAAGACTTCAATTTTAATTTGGTGATGGCACAGGTATGGATATGCGCT	426
Db	1389	TT	GTCTCACATTAACAGATACCAAGCTTCTTGATTTGGAGATGGCACAGGTCGTGGTTCGGCC	1330
QY	427	GAA	AGTCATTAATTTTCGCTTTTGGGCATGTGAATATAGACGTCACTTTTAAAGTTATATAA	486
Db	1329	AA	TGCCAAATATTTTGTCTTTGAATCTGACGAATATGACGCTCACTTCATGCTTTACCAAC	1270
QY	487	CT	GAATTCACGCAATATATGACAAAAATTTGATTTTCCGATCATCTCTGATATTTTAAAGATAT	546
Db	1269	CC	AGAACTACTTATTAACCAACATTCGACTTTGACCATCCAGATATTTTCAAGTCTC	1210
QY	547	AA	TGATGTTTTTGGATGCAATTCCAAGAAATGGCACATATGTTTAAAAAGGTATTTATGCT	606
Db	1209	GAG	ATGTTTTTAAATGCTTTTAACGACTATGCCAAACAAATCACCAAGGTCCTTTTGTGTC	1150
QY	607	TG	GGGTGATGATGAACATCTACGTAAAAATTTGAACGAGATGTTTCAATTTATTTATGGA	666
Db	1149	TAT	GGTGAAGATGCTGAATTTGCGTARGATTAACGCTGTGATGACCAAAATTTATTTATGCT	1090
QY	667	TTT	AAAAATTCGG---ATGACATTTATGCTCAAAATATTTCAAAATTTACGATAAAGTACT	723
Db	1089	TTT	GAGCTGAAGCAATGACTTTGTAGCTAGTGATCTTCTCGTTCAATACTGCTTCA	1030
QY	724	GC	TTTGTATGTGTATGPGGATGGTGATTTTATGATCACTTCCTGTCTCCACAATATGCT	783
Db	1029	AC	CTTCACCGTTCAATTTCCGTGACAAAACTTGGGGCAATTTCCACATTTCCAACTTTGCT	970
QY	784	GAC	CATACAGTTTAAATGCAATTAGCTGTAAATTTGCGATTTAGTTATTTAGAGAAGCTAGAT	843
Db	969	CG	TCAAAATATCATGATGACGACGCGTTTATTGCTCTTCTTTACACAGCAGGATTTGAT	910
QY	844	GT	TACAAATATTAAGAAGCATTAGAAAAAGTTTGGTGGTGTAAACGTCGTTTCAATGAA	903
Db	909	TT	GAACTTGGTGGTGAGCACTTGAACAAATTTGCCCGGTGTAAACGTCGTTTCACTGAG	850
QY	904	ACT	CAATTTGCAAAATCAAGTTATTTAGATGATTTATGCACACCATCCAAAGAGAATTAGT	963
Db	849	AAA	ATTGTCAATGATACAGTGATTAATCGATGACTTTTGCCCAACCATCCACAGAAATTAAT	790
QY	964	GCT	CAATTTGAAACGACGACGAAGAATATCCACATAAAGAAGTTGTTGCAATTTCAA	1023
Db	789	GG	ACTTGGATGCGGCTCGTCAGAAATATACCAAGCAAGGAAATTTGAGCAGTCTTTCAA	730
QY	1024	CC	ACACACTTTCTCTAGAACACAGGCATTTTAAATGAAATTTCCAGAAAGTTTAAAGTAA	1083
Db	729	CG	CAATCCTTTACAGAACCAATGTCCTTTGTGGACGACTTTGCCCAATGCTTTAAACCAA	670
QY	1084	GC	AGATCGGTATTTCTTATGTGAAATTTTTTGGATCAATTAG---AGAAAAATCTGGCGCA	1140
Db	669	GC	AGATGCTGTTTATCTAGCGCAATTTATGGCTCGGCTCGTGAAGTAGATCATGGTGAC	610
QY	1141	TT	ACGATACAGATTTTAATGATAAAT---TGAAGGTGCATCGTTAATTAATGAAGAT	1197
Db	609	GT	TAGGTAGAGACCTAGCCCAACAAATTCACAAAAAACAACCAAGTGATTACTGTTGAA	550
QY	1198	TCT	ATTAATGATTAGAACAAATTTGATATGCTGTTATTTTATTTATGGTGCAGGTAT	1257
Db	549	AA	TGTTTCTCACTCCTAGAACCATGACAAATGCTGTTTACGTCTTTATGGAGCAGGAGAC	490
QY	1258	ATT	CAAAATTTACAAAATGCAATAT	1281

```
Db      489 ATCCAAACCTATGAATACCTATTT 466
|||||
RESULT 5
US-08-940-572-1
; Sequence 1, Application US/08940572
; Patent No. 6310193
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; APPLICANT: Black, Michael T.
; APPLICANT: Hodgson, John E.
; APPLICANT: Knowles, David J.
; APPLICANT: Lonetto, Michael A.
; APPLICANT: Nicholas, Richard O.
; APPLICANT: Stodola, Robert K.
; TITLE OF INVENTION: No. 6310193el Murc
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/940,572
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024022
; FILING DATE: 16-AUG-1996
; APPLICATION NUMBER: US 08/899711
; FILING DATE: 08-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Q. Todd
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50533-04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215/994-2252
; TELEFAX: 215/994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2715 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-940-572-1
Query Match      30.3%; Score 399.2; DB 4; Length 2715;
Best Local Similarity 59.1%; Pred. No. 1.4e-83;
Matches 759; Conservative 0; Mismatches 513; Indels 12; Gaps 4;

Qy      10 TATCATTTTTCGGAAATTAAAGGTTCTGGCATGAGTTCAATTAGCACAAATCATGCATGAT 69
Db      408 TATCATTTTATCGGAATTAAAGGATCAGGATGAGTGCCTTGGCTTGAATGTTGCACCA 467

Qy      70 TTAGACATGAGTTCAAGGATCGGATATTGAGAACTAGTATTACAGAAAGTTGCTCTTT 129
Db      468 ATGGGGCAACAGGTTTCAGGGATCAGATGTTGAAAAGTACTACTTTACCCCAACGCGGTCTT 527

Qy      130 AGAAATAGGGGATAAAAATATTACATTTGATGCTTAATAACATAAAGAAAGATAGTGA 189
Db      528 GAGCAGGAGGAATTTACATCTTCTTTTGAAGAAAGATCTAGACGGTATATGGAA 587

Qy      190 GTTATACAGGTAATGCAATT---CGCGAGTAGCCCATGAGAAATAGTACGTGCACATCAA 246
|||||
```

Db 588 ATTATCGCTGGAAATGCCCTTTTCGTCCAGATAACAACGTGCAAAATGCTTANGCGGACCAA 647

Qy 247 TTGAAATTAGATGTTGTAAAGTTATATGATATTTTATAGGACAGATTTATGATCAATATACT 306

Db 648 AATGGTATCAGCTACAAACGTTACCATGAGTTTCTAGGTAGCTTTATGCGTGACTTTGTT 707

Qy 307 TCAGTAGCTGTAACTGGTGCAATGGTAAATCTTCAACAACAGGTTTATTTATCACATGTT 366

Db 708 AGCATGGGAGTAGCAGGAGCAATCGAAAAAATCTCAACGACAGGTATGTTGTTCTCATGTC 767

Qy 367 ATGATGCTGTATAAAAAGACACTTCATTTTAATTTGGTGATGGCACAGCATATGGGATGCGCT 426

Db 768 TTGTCTCATAATTACAGATACCAAGCTTCTTGATTTGGAGATGGGACAGGTCGTGCTCGGCC 827

Qy 427 GAAAGTGATTTATTTTCGCTTTTGGAGCATGTGTAATATAGACGTCACCTTTTAAAGCTATATAA 486

Db 828 AATGCCAATATTTTGTCTTTTGAATCTGACGAATATGAGCGTCACCTTCATGCTTTACCA 887

Qy 487 CCTGATTAACGCAATTTATGACAAATATGATTTTCGATCATCTCTGATTTTAAAGATATT 546

Db 888 CCAGAACTACTCTATTATCAACCAATTGACTTTTGACCATCCAGATTTATTTTCAACAGTCTC 947

Qy 547 AATGATGTTTTTGTGATTCCTCAAGAAATGGCACATAATGTTTAAAAAAGGTATATTGCT 606

Db 948 GAGGATGTTTTCAATGCTTTTAAACGACTATGCAAAACAAATTTACCAAGGGTCTTTTGTG 1007

Qy 607 TGGGCTGATGATGAACATCTAGTAAATTTGAAGCAGATGTTTCCAAATTTATTTATTTATGGA 666

Db 1008 TATGCTGAAGATGCTGAATGCGTGAAGATTAGCTCTGATGCAACCAATTTATTTATTTATGCT 1067

Qy 667 TTTAAAGATTCGG---ATGACATTTTATGCTCAAAATATTCAAATTTACGATTAAGGTACT 723

Db 1068 TTTGAAGCTGAAGCAATGACTTTTGTAGCTAGTGATCTTCTCGTTCAACACTGCTGTTCA 1127

Qy 724 GCTTTGATGTTGATGGATGGTGGAGTTTATGATCACTTCTCTGCTCCCAAAATGCT 783

Db 1128 ACCTTCCCGTTCAATTTCCGTGGACAAAACTTGGGGCAATTTCCCATTTCCAACTTTGCT 1187

Qy 784 GACCATACAGTTTAAATGCAATAGCTGTAATGCGATTAGTTTATTTAGAGAAGCTAGAT 843

Db 1188 CGTCACAAATATCATGAATGCGACAGCCGTTATTTGCTCTTTTACACAGCAGGATTTGAT 1247

Qy 844 GTTACAAATATTAAAGAACGATTAGAAACGTTTGGTGGTGTATAACGTCGTTTCAATGAA 903

Db 1248 TTGAACCTTGTGCGTGAGCACTTGAAACATTTGCCGCTGTATAACGCTGTTTCACTGAG 1307

Qy 904 ACTACAAATGCAATCAAGTTATTGTAGATGATTATGACACCATATGACACCAAGAAATAGT 963

Db 1308 AAAATTGTCATGATACAGTGATTTATGATGACTTTGGCCCATATCCCAACAGAAATATT 1367

Qy 964 GCTACAAATGAAACAGCACGAAAGAAATATCCATAAAGAAAGTTGTTGCGAGTATTTCAT 1023

Db 1368 GCGACCTTGGATGCGGCTCGTCAGAAATATCCCAAGCAAGGAAATTTGTAGCAGTCTTTCAA 1427

Qy 1024 CCACACACTTTTCTTAGAACACAGGCAATTTTAAATGAATTTGCAAGAAATTTTAAGTAAA 1083

Db 1428 CCGATACCTTTTCAAGAACCAATTTGCCCTGTTGGACGACTTTGGCCATGCTTTTAAACCAA 1487

Qy 1084 GCAGATCGTGTATTCTTATGTAATTTTGGATCAATTTAG---AGAAAAATACTCGCGCA 1140

Db 1488 GCAGATGCTGTTTATCTAGCGCAATTTTATGGCTCGGCTCGTGAAGTAGATCATGCTGAC 1547

Qy 1141 TTAACGATACAGATTTAAATTTGATAAAAT---TGAAGGTGCATCGTTAAATTAATGAAGAT 1197

Db 1548 GTTAAAGGTAGAAGACCTTAGCCAAATAAAATCAACAAAAACCAACCAAGTATTACTGTTGAA 1607

Qy 1198 TCTATTAAATGTTATAGAACAAATTTGATAATGCTGTTATTTTATTTATGGGTGCGAGTAT 1257

Db 1608 AATGTTTCTCCATCTCTAGACCATGACATGCTGTTTACGCTTTTATGGGAGCAGGAGAC 1667

Qy 1258 ATTCAAAAATTTACAAAATGCATAT 1281

Db 1668 ATCCAAACCTATGAAATCTCAATTT 1691

RESULT 6
US-08-714-918-61/c
; Sequence 61, Application US/08714918
; Patent No. 6037123
; GENERAL INFORMATION:
; APPLICANT: Benton, Bret
; APPLICANT: Lee, Ving
; APPLICANT: Malouin, Francois
; APPLICANT: Martin, Patrick K.
; APPLICANT: Schmid, Molly B.
; APPLICANT: Sun, Dongxu
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
; TITLE OF INVENTION: TARGET GENES
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,918
; FILING DATE: September 13, 1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,102
; FILING DATE: December 22, 1995
; APPLICATION NUMBER: 60/003,798
; FILING DATE: September 15, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 222/005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 888 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-714-918-61

Query Match 29.6%; Score 389.6; DB 3; Length 888;
Best Local Similarity 94.1%; Pred. No. 1.7e-81;
Matches 462; Conservative 1; Mismatches 21; Indels 7; Gaps 6;
QY 828 TTAGAGAAGCTAGATTGTTACAATATTAAGAAGCATTAGAAACGTTTGGTGGTTAA 887
DB 888 TTAGAGAAGTTAGATG-TCCCAATNTTAAAGAAGCATTAGAAACG-TTGGTGGTTAA 831
QY 888 AGCTGCTTTCAATGAACACTCAATTCGAATCAAGTTATTGTAGATGATTATGCACCA 947
DB 830 AGCTC-NNTCNATGANACTCAATCGCAATCAAGTTAATGTAGCTGATTATGNACCA 772
QY 948 TCAAGAGAAATTAGTGCTACAAATGAAACAGCAGCAAGAAATATCCATTAAGAGT 1007
DB 771 TCAAGAGAAATTAGTGCTGNNCAATTAAGAAACCGCAAGAAATATCCATTAAGAGT 712
QY 1008 TGTGTCAGTATTTTC-AACCACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTG 1066
|||||

DB 711 TGTTCAGTATNTCAAACCAACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTG 652
QY 1067 CAGAAAGTTTAAAGTAAAGCAGATCGTGATT--CTTATGTGAAATTTTGGATCAATTAG 1124
DB 651 CAGAAAGTTTAAAGTAAAGCAGATCGTGATTCTTATTGTGAAATTTTGGATCAATTAG 592
QY 1125 AGAAATACTGGCGCATTAACGATACAAGATTAAATTGATAPAAAATTGAAGGTGCATCGTT 1184
DB 591 AGAAATACTGGCGCATTAACGATACAAGATTAAATTGATAPAAAATTGAAGGTGCATCGTT 532
QY 1185 AATTAATGAAGATTCCTATTAAATGATAGAACAAATTTGATAGTCTGTTATTTT-ATTTA 1243
DB 531 AATTAATGAAGATTCCTATTAAATGATAGAACAAATTTGATAGTCTGTTATTTTCAATTA 472
QY 1244 TGGGTGCAGTGATATTTCAAAATTAACAAATGCATATTTAGATAAATTAGGCATGAAA 1303
DB 471 TGGGTGCAGTGATATTTCAAAATTAACAAATGCATATTTAGATAAATTAGGCATGAAA 412
QY 1304 ATGCGTTTTAA 1314
DB 411 ATGCGTTTTAA 401

RESULT 7
US-09-265-315-61/c
; Sequence 61, Application US/09265315
; Patent No. 6187541
; GENERAL INFORMATION:
; APPLICANT: Benton, Bret
; APPLICANT: Lee, Ving J.
; APPLICANT: Malouin, Francois
; APPLICANT: Martin, Patrick K.
; APPLICANT: Schmid, Molly B.
; APPLICANT: Sun, Dongxu
; TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
; TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,315
; FILING DATE: March 9, 1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/714,918
; FILING DATE: September 13, 1996
; APPLICATION NUMBER: 60/009,102
; FILING DATE: December 22, 1995
; APPLICATION NUMBER: 60/003,798
; FILING DATE: September 15, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 240/247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:


```
; LENGTH: 888 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-265-315-61

Query Match      29.6%; Score 389.6; DB 3; Length 888;
Best Local Similarity 94.1%; Pred. No. 1.7e-81;
Matches 462; Conservative 1; Mismatches 21; Indels 7; Gaps 6;

QY 828 TTTAGAGAAGCTAGATGTTTACAAATATTAAGAAGCATTAGAAACGCTTTGGTGGTGTAA 887
Db 888 TTTAGAGAAGTTAGTG-TCCCAATNTTAAAGAAGCATTAGAAACG-TTGGTGGNGTTAA 831

QY 888 ACGTCTGTTTCAATGAACTACAAATGCAATCAAGTTATTGTAGATGATTATGCACACCA 947
Db 830 ACGTC-NNTCNATGANACTACAAATGCAATCAAGTTATTGTAGATGATTATGNACACCA 772

QY 948 TCCAAAGAGAAATTAGTGCTACAAATGAAACAGCAGAAAGAAATATCCACATAAAGAAGT 1007
Db 771 TCCAAAGAGAAATTAGTGCTACAAATGAAACAGCAGAAAGAAATATCCACATAAAGAAGT 712

QY 1008 TGTTCGAGTATTC-AACACACACCTTCTCTAGAACACAGGCATTTTAAATGATTTG 1066
Db 711 TGTTCGAGTATNTCAACACACACACTTTCTCTAGAACACAGGCATTTTAAATGATTTG 652

QY 1067 CAGAAAGTTTAAAGTAAAGCAGATCGTGATT--CTTATGTGAAATTTTGGATCAATTAG 1124
Db 651 CAGAAAGTTTAAAGTAAAGCAGATCGTGATT--CTTATGTGAAATTTTGGATCAATTAG 592

QY 1125 AGAAATACTGGCGCAATTAACGATCAAGATTTAAATTTGATAAAATTCGAAGTGCATCGTT 1184
Db 591 AGAAATACTGGCGCAATTAACGATCAAGATTTAAATTTGATAAAATTCGAAGTGCATCGTT 532

QY 1185 AATTAATGAAGATTCCTATTAAATGATTAGAACAAATTTGATAAATTCGCTGTTATTTT-ATTTA 1243
Db 531 AATTAATGAAGATTCCTATTAAATGATTAGAACAAATTTGATAAATTCGCTGTTATTTTCAATTTA 472

QY 1244 TGGGTGCAGGTGATATTCAAAATTAACAAAATTCATAATTTAGATAAAATTAGGCATGAAA 1303
Db 471 TGGGTGCAGGTGATATTCAAAATTAACAAAATTCATAATTTAGATAAAATTAGGCATGAAA 412

QY 1304 ATGCGTTTTAA 1314
Db 411 ATGCGTTTTAA 401

RESULT 8
US-09-265-315-61/c
; Sequence 61, Application US/09265315
; Patent No. 6187541
; GENERAL INFORMATION:
; APPLICANT: Benton, Bret
; APPLICANT: Lee, Ving J.
; APPLICANT: Malouin, Francois
; APPLICANT: Martin, Patrick K.
; APPLICANT: Schmid, Molly B.
; APPLICANT: Sun, Dongxu
; TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
; TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
; TITLE OF INVENTION: TARGET GENES
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
```

```
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,315
; FILING DATE: March 9, 1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/714,918
; FILING DATE: September 13, 1996
; APPLICATION NUMBER: 60/009,102
; FILING DATE: December 22, 1995
; APPLICATION NUMBER: 60/003,798
; FILING DATE: September 15, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 240/247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 888 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-265-315-61

Query Match      29.6%; Score 389.6; DB 3; Length 888;
Best Local Similarity 94.1%; Pred. No. 1.7e-81;
Matches 462; Conservative 1; Mismatches 21; Indels 7; Gaps 6;

QY 828 TTTAGAGAAGCTAGATGTTTACAAATATTAAGAAGCATTAGAAACGCTTTGGTGGTGTAA 887
Db 888 TTTAGAGAAGTTAGTG-TCCCAATNTTAAAGAAGCATTAGAAACG-TTGGTGGNGTTAA 831

QY 888 ACGTCTGTTTCAATGAACTACAAATGCAATCAAGTTATTGTAGATGATTATGCACACCA 947
Db 830 ACGTC-NNTCNATGANACTACAAATGCAATCAAGTTATTGTAGATGATTATGNACACCA 772

QY 948 TCCAAAGAGAAATTAGTGCTACAAATGAAACAGCAGAAAGAAATATCCACATAAAGAAGT 1007
Db 771 TCCAAAGAGAAATTAGTGCTACAAATGAAACAGCAGAAAGAAATATCCACATAAAGAAGT 712

QY 1008 TGTTCGAGTATTC-AACACACACCTTCTCTAGAACACAGGCATTTTAAATGATTTG 1066
Db 711 TGTTCGAGTATNTCAACACACACACTTTCTCTAGAACACAGGCATTTTAAATGATTTG 652

QY 1067 CAGAAAGTTTAAAGTAAAGCAGATCGTGATT--CTTATGTGAAATTTTGGATCAATTAG 1124
Db 651 CAGAAAGTTTAAAGTAAAGCAGATCGTGATT--CTTATGTGAAATTTTGGATCAATTAG 592

QY 1125 AGAAATACTGGCGCAATTAACGATCAAGATTTAAATTTGATAAAATTCGAAGTGCATCGTT 1184
Db 591 AGAAATACTGGCGCAATTAACGATCAAGATTTAAATTTGATAAAATTCGAAGTGCATCGTT 532

QY 1185 AATTAATGAAGATTCCTATTAAATGATTAGAACAAATTTGATAAATTCGCTGTTATTTT-ATTTA 1243
Db 531 AATTAATGAAGATTCCTATTAAATGATTAGAACAAATTTGATAAATTCGCTGTTATTTTCAATTTA 472

QY 1244 TGGGTGCAGGTGATATTCAAAATTAACAAAATTCATAATTTAGATAAAATTAGGCATGAAA 1303
Db 471 TGGGTGCAGGTGATATTCAAAATTAACAAAATTCATAATTTAGATAAAATTAGGCATGAAA 412

QY 1304 ATGCGTTTTAA 1314
Db 411 ATGCGTTTTAA 401

RESULT 9
US-09-266-417-61/c
```

Sequence 61, Application US/09266417

Patent No. 6228588
GENERAL INFORMATION:
APPLICANT: Benton, Bret
APPLICANT: Lee, Ving J.
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
TITLE OF INVENTION: TARGET GENES
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/266,417
FILING DATE: March 9, 1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/714,918
FILING DATE: September 13, 1996
APPLICATION NUMBER: 60/009,102
FILING DATE: December 22, 1995
APPLICATION NUMBER: 60/003,798
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 240/248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 61:
LENGTH: 888 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-266-417-61

Query Match 29.6%; Score 389.6; DB 3; Length 888;
Best Local Similarity 94.1%; Pred. No. 1.7e-81;
Matches 462; Conservative 1; Mismatches 21; Indels 7; Gaps 6;

QY 828 TTTAGAGAACTAGATTGCTTACAAATTAAGAAGCATTAGAAAGTTGGTGGTGTAA 887
Db 888 TTTAGAGAACTTAGATG-TCCCAATNTTAAAGAAGCATTAGAAAGC-TTGGTGGGTAA 831
QY 888 ACCTCGTTTCATGAACTACAATGCAATCAAGTTATTCTAGATGATTATGCACCA 947
Db 830 ACCTC-NNTCNATGANNACTACATCGCAATCAAGTTAATAGTCTGATTAGNACCA 772
QY 948 TCCAAGAGAAATTAGTCTTACAAATGAAACAGCAGCAAGAAATATCCACATAAGAGT 1007
Db 771 TCCAAGAGAAATTAGTGCNNCAATGAAACCGCAGCAAGAAATATCCACATAAGAGT 712
QY 1008 TGTTGAGTATTC-AACACACACTTCTCTAGAACACAGGCATTTTAAATGAATTG 1066
Db 711 TGTTGAGTATNTCAACACACACTTCTCTAGAACACAGGCATTTTAAATGAATTG 652

QY 1067 CAGAAAGTTTAAAGTAAAGCAGATCGTGTATT--CTTATGTGAAATTTTGGATCAATTAG 1124
Db 651 CAGAAAGTTTAAAGTAAAGCAGATCGTGTATTCTTATTGTGAAATTTTGGATCAATTAG 592
QY 1125 AGAAATATCTGGCGCATTAACGATACAGATTTAATTGATAAAATGGAAGTGCATGTT 1184
Db 591 AGAAATATCTGGCGCATTAACGAWRCAAGATTTAATTGATAAAATGGAAGTGCATGTT 532
QY 1185 AATTAATGAAGATTCATTAAATCTATTAGAACAAATTTAGATAATTCCTGTTATTTT-ATTTA 1243
Db 531 AATTAATGAAGATTCATTAAATCTATTAGAACAAATTTAGATAATTCCTGTTATTTTCAITTA 472
QY 1244 TGGTGCAGGTGATATTCAAAATTTACAAATTCATATTTAGATAAATTAGGCATGAAA 1303
Db 471 TGGTGCAGGTGATATTCAAAATTTACAAATTCATATTTAGATAAATTAGGCATGAAA 412
QY 1304 ATGCGTTTTAA 1314
Db 411 ATGCGTTTTAA 401

RESULT 10

US-09-528-709-61/c
Sequence 61, Application US/09528709
Patent No. 6630303

GENERAL INFORMATION:

APPLICANT: Benton, Bret
Lee, Ving
Malouin, Francois
Martin, Patrick K.
Schmid, Molly B.
Sun, Dongxu

TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL

NUMBER OF SEQUENCES: 111

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
Storage

COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/528,709
FILING DATE: 17-Mar-2000
CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/714,918
FILING DATE: September 13, 1996
APPLICATION NUMBER: 60/009,102

FILING DATE: December 22, 1995

APPLICATION NUMBER: 60/003,798

FILING DATE: September 15, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 222/005

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 61:

SEQUENCE CHARACTERISTICS:

LENGTH: 888 base pairs

TYPE: nucleic acid

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-09-528-709-61

Query Match          29.6%; Score 389.6; DB 4; Length 888;
Best Local Similarity 94.1%; Pred. No. 1.7e-81;
Matches 462; Conservative 1; Mismatches 21; Indels 7; Gaps 6;

QY 828 TTATGAGAGCTAGATTGTTACAAATATTAAAGAGCAATTAGAAACGTTTGGTGGTTAA 887
Db      |||||
QY 888 TTATGAGAGCTAGATTGTTACAAATNTTAAAGAGCAATTAGAAACG-TTGGTGGTTAA 831
Db      |||||
QY 888 ACCTCGTTTCAATGAAACTACAAATTCGAATTCGAATTTGTTAGATGATTATGCACCA 947
Db      |||||
QY 830 ACCTC-NNTCNATGANACTACAAATCGCAATCAAGTTAATGTAGCTGATTATGNACCA 772
Db      |||||
QY 948 TCCAGAGAGAAATTAGTGTACAAATTCGAACAGCAGCAAGAAATATCCACATAAGAGT 1007
Db      |||||
QY 771 TCCAGAGAGAAATTAGTGTACAAATTCGAACAGCAGCAAGAAATATCCACATAAGAGT 712
Db      |||||
QY 1008 TGTTCAGTATTC-AACACACACATTTCTCTAGAACACAGGCAATTTTAAATGAATTTG 1066
Db      |||||
QY 711 TGTTCAGTATNTCAACACACACATTTCTCTAGAACACAGGCAATTTTAAATGAATTTG 652
Db      |||||
QY 1067 CAGAAAGTTTAAAGTAAAGCAGATCGTGTATT--CTTATGTGAATTTTGGATCAATTAG 1124
Db      |||||
QY 651 CAGAAAGTTTAAAGTAAAGCAGATCGTGTATTCTTATTGTGAATTTTGGATCAATTAG 592
Db      |||||
QY 1125 AGAAATATCTGCGCATTAACGATACAAATTCGAACAGCAGCAAGAAATATCCACATAAGAGT 1184
Db      |||||
QY 591 AGAAATATCTGCGCATTAACGATACAAATTCGAACAGCAGCAAGAAATATCCACATAAGAGT 532
Db      |||||
QY 1185 AATTAATGAAGATTTCTATTAAATGTTATAGAACAAATTTGATAATGCTGTTATTTT-ATTTA 1243
Db      |||||
QY 531 AATTAATGAAGATTTCTATTAAATGTTATAGAACAAATTTGATAATGCTGTTATTTTCAATTA 472
Db      |||||
QY 1244 TGGTGCAGGTGATATTCACAAATTTACAAATTCGATATTTAGATAAATTAGGCATGAAA 1303
Db      |||||
QY 471 TGGTGCAGGTGATATTCACAAATTTACAAATTCGATATTTAGATAAATTAGGCATGAAA 412
Db      |||||
QY 1304 ATGCGTTTAA 1314
Db      |||||
QY 411 ATGCGTTTAA 401
Db      |||||

RESULT 11
US-09-527-745-61/c
; Sequence 61, Application US/09527745
; Patent No. 6638718
; GENERAL INFORMATION:
; APPLICANT: Benton, Bret
; Lee, Ving
; Malouin, Francois
; Martin, Patrick K.
; Schmid, Molly B.
; Sun, Dongxu
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
; TARGET GENES
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
```

```
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/527,745
; FILING DATE: 17-Mar-2000
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,918
; FILING DATE: September 13, 1996
; APPLICATION NUMBER: 60/009,102
; FILING DATE: December 22, 1995
; APPLICATION NUMBER: 60/003,798
; FILING DATE: September 15, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 222/005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 888 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-09-527-745-61
```

```
Query Match          29.6%; Score 389.6; DB 4; Length 888;
Best Local Similarity 94.1%; Pred. No. 1.7e-81;
Matches 462; Conservative 1; Mismatches 21; Indels 7; Gaps 6;

QY 828 TTATGAGAGCTAGATTGTTACAAATATAAGAGCAATTAGAAACGTTTGGTGGTTAA 887
Db      |||||
QY 888 ACCTCGTTTCAATGAAACTACAAATTCGAATTCGAATTTGTTAGATGATTATGCACCA 947
Db      |||||
QY 830 ACCTC-NNTCNATGANACTACAAATCGCAATCAAGTTAATGTAGCTGATTATGNACCA 772
Db      |||||
QY 948 TCCAGAGAGAAATTAGTGTACAAATTCGAACAGCAGCAAGAAATATCCACATAAGAGT 1007
Db      |||||
QY 771 TCCAGAGAGAAATTAGTGTACAAATTCGAACAGCAGCAAGAAATATCCACATAAGAGT 712
Db      |||||
QY 1008 TGTTCAGTATTC-AACACACACATTTCTCTAGAACACAGGCAATTTTAAATGAATTTG 1066
Db      |||||
QY 711 TGTTCAGTATNTCAACACACACATTTCTCTAGAACACAGGCAATTTTAAATGAATTTG 652
Db      |||||
QY 1067 CAGAAAGTTTAAAGTAAAGCAGATCGTGTATT--CTTATGTGAATTTTGGATCAATTAG 1124
Db      |||||
QY 651 CAGAAAGTTTAAAGTAAAGCAGATCGTGTATTCTTATTGTGAATTTTGGATCAATTAG 592
Db      |||||
QY 1125 AGAAATATCTGCGCATTAACGATACAAATTCGAACAGCAGCAAGAAATTCGAAGTGCATCGT 1184
Db      |||||
QY 591 AGAAATATCTGCGCATTAACGATACAAATTCGAACAGCAGCAAGAAATTCGAAGTGCATCGT 532
Db      |||||
QY 1185 AATTAATGAAGATTTCTATTAAATGTTATAGAACAAATTTGATAATGCTGTTATTTT-ATTTA 1243
Db      |||||
QY 531 AATTAATGAAGATTTCTATTAAATGTTATAGAACAAATTTGATAATGCTGTTATTTTCAATTA 472
Db      |||||
QY 1244 TGGTGCAGGTGATATTCACAAATTTACAAATTCGATATTTAGATAAATTAGGCATGAAA 1303
Db      |||||
QY 471 TGGTGCAGGTGATATTCACAAATTTACAAATTCGATATTTAGATAAATTAGGCATGAAA 412
Db      |||||
QY 1304 ATGCGTTTAA 1314
Db      |||||
QY 411 ATGCGTTTAA 401
Db      |||||
```

```
RESULT 12
US-08-961-083-115
; Sequence 115, Application US/08961083
```


OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/536,784
FILING DATE: 30-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: OCT-30-1997
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB340P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 115:
SEQUENCE CHARACTERISTICS:
LENGTH: 1267 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 115:
US-09-536-784-115

Query Match 28.1%; Score 370.2; DB 4; Length 1267;
Best Local Similarity 58.6%; Pred. No. 5.9e-77;
Matches 721; Conservative 0; Mismatches 498; Indels 12; Gaps 4;
QY 63 GCATGATTAGACATGAAGTTCAAGGATCGGATTTGGAAGTACGTTATTTACAGAGT 122
DB 1 GCACAGATGGGCAACAGGTTTCAGGATCAGATGTTGAAAAGTACTACTTTACCCACG 60
QY 123 TGCCTTTAGAAATAGGGGATAAAATATTACATTTGATGCTTAATAACATAAAAGAGA 182
DB 61 CGGCTTGAGCAGGAGGAAATACCATTTCTCTTTTGTATGAAAAAATCTAGACGGTA 120
QY 183 TATGGTAGTTATACAGGTAATGCAATT---CGCGAGTAGCCATGAGAAATAGTAGTGTC 239
DB 121 TATGGAATATTGCTGGAAATGCTTTCGTCAGATACACGTCGAAATGTCCTATGC 180
QY 240 ACATCAATTAATGATAGTTGTAGTTATTAATGATTTTATTTAGGACAGATTATGATCA 299
DB 181 GGACCAAAATGGTATACAGTACAAACGTTACCATGAGTTTCTAGTAGCTTTATGGTGA 240
QY 300 ATATACCTCAGTAGCTGTAACCTGGTGACATGTAAGCTTCTACAAAGGTTTATATC 359
DB 241 CTTTGTAGCATGGAGTAGGAGGACATGGAATACTTCAACGACAGGTATGTTGTC 300
QY 360 ACATGTTATGAATGGTGATAAAAGACTTTCATTTTAAATTTGATGGCACAGGTATGGG 419
DB 301 TCATGCTCTGCTCACATTACAGATACCAAGCTTCTTGATTTGAGATGGACAGGTCTGG 360
QY 420 ATTGCTGAAAGTATGATTTTTCGCTTTTGGAGCATGTAATATAGACGTCATTTTAA 479
DB 361 TTCGGCCAAATCCAAATATTTGCTTTTGAATCTGACGAATATGAGCGTCACTTCATGCC 420
QY 480 TTATAAACCTGATACGCAATATGACAAATATTGATTTTCGATCATCTGATATTATTA 539
DB 421 TTACCCACCAAGTACTCTATATACCAACATTCATTTGACCATCCAGATTATTTTAC 480
QY 540 AGATATTAAATGATGTTTGTGATGCTTCCAGAAATGGCACATAATGTTTAAAAAGGTAT 599
DB 481 AAGTCTCGAGATGTTTAAATGCTTTTAAAGCTATGCGCAAAATACCAAGGGTCT 540
QY 600 TATTGCTTGGGTGATGATGAACATCTACGTAATAATTTGAAGCAGATGTTTCCAAATTA 659
DB 541 TTTTGTCTATGTTGAAGATGCTGAATTCGTAAGATTAGCTCTGATGACCAATTTATTA 600
QY 660 TTATGGATTAAAGATTCGG---ATGACATTTATGCTCAAAATATTCAAAATACGATTA 716
DB 601 TTATGGTTTGAAGCTGAAGCAATGACTTTGTAGCTAGTATCTTCTTCTTCAATAC 660

QY 717 AGGTACTGCTTTTGTATGTTAGTGGATGGTGGAGTTTATGATCACTTCTCTCTCCACA 776
DB 661 TGGTTCAACCTTTCACCGTTCAATTTCCGTCGACAAAACCTTGGGCAATTTCCACATTTCAAC 720
QY 777 ATATGTTGACCATACAGTTTAAATGATTAAGTCTTAATTCGATTAATTCGATTAATTAAGAA 836
DB 721 CTTTGGTCTGTCACAAATATCATGAATGCGACGCGTTATTTGGTCTTCTTTACACAGCAG 780
QY 837 GCTAGATGTTTCAAAATATTTAAAGAACGATTTAGAAACGTTTGGTGTGTTTAAACGTCGTTT 896
DB 781 ATTGATTTGACTTGGTGGTGGACCTTGAAACATTTGCCGCTGTTTAAACGTCGTTT 840
QY 897 CAATGAAACTTACAAATTCGAAATCAAGTTATTTGATGATGATTTATGACACCATCCAGAGA 956
DB 841 CACTGAGAAAATTTGCTCAATGATACAGTATTCGATGACTTTTGCCCACTCCCAACAGA 900
QY 957 AATTAGTCTACAAATTTGAAACAGACGACGAAAGAAATATCCATAAAGATTTGTTGCAGT 1016
DB 901 AATTATTTGCGACCTTGGATGCGGCTCGTCAGANAATCCCAAGCAAGGAAATTTGTAGCAGT 960
QY 1017 ATTTCAACACACACTTTTCTAGAACACAGGCAATTTTAAATGAATTTGCAGAAAGTTT 1076
DB 961 CTTTCAACCGCATACCTTTTACAGAACCAATTTGCGCTCGTGGAGTATGATCA 1080
QY 1077 AAGTAAAGCAGATCGTGTATTTCTTATGTGAATTTTGGATCAATTAG---AGAAAATAC 1133
DB 1021 AAACCAAGCAGATGCTGTTTATCTAGCGAAATTTATGGCTCGTGGAGTATGATCA 1080
QY 1134 TGGCGCATTAACGATACAGATTTAATGTATAAAT---TGAAGTGATCGTTAATTA 1190
DB 1081 TGGTACGTTTAAAGGTAGACACCTAGCCAAACAAATCAACAAAAACACCAAGTGATTAC 1140
QY 1191 TGAAGATTTCTTAAATGATTTAGAACAAATTTGATAATGCTGTTATTTTATTTATGGTGC 1250
DB 1141 TGTGAAATGTTTCTCCACTCTAGACCATGACATGCTGTTTACGTTTATGGGAGC 1200
QY 1251 AGGTGATTTCAAAATTTACAAAATGCAAT 1281
DB 1201 AGGAGATCCAAACCTATGATATCAATTT 1231

RESULT 14

US-09-107-532A-3260
; Sequence 3260, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS: 7310
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489

```

; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3260:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 873 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...873
; SEQUENCE DESCRIPTION: SEQ ID NO: 3260:
US-09-107-532A-3260

Query Match      26.9%; Score 354.4; DB 4; Length 873;
Best Local Similarity 64.0%; Pred. No. 2.4e-73;
Matches 535; Conservative 0; Mismatches 301; Indels 0; Gaps 0;

QY 10 TATCATTTTTCGGAATTAAGGTTCTGGCAGTGGTTCATTTAGCAGCAATCATCATGAT 69
DB 10 TATCATTTTTCGGAATTAAGGTTCTGGCAGTGGTTCATTTAGCAGCAATCATCATGAT 69
QY 37 TATCATTTTTCGGAATTAAGGTTCTGGCAGTGGTTCATTTAGCAGCAATCATCATGAT 96
DB 37 TATCATTTTTCGGAATTAAGGTTCTGGCAGTGGTTCATTTAGCAGCAATCATCATGAT 96
QY 70 TTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTTATTTACAGAGTTGCTTT 129
DB 70 TTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTTATTTACAGAGTTGCTTT 129
QY 97 CRAAGTCTAAATGTCGAGGATTCAGACATTTGAAATATTTCTTTACAAAGAGATTTA 156
DB 97 CRAAGTCTAAATGTCGAGGATTCAGACATTTGAAATATTTCTTTACAAAGAGATTTA 156
QY 130 AGAATAAGGGGATAAATAATTACCATTTGATCTAATAACATAAAGAAAGATATGTA 189
DB 130 AGAATAAGGGGATAAATAATTACCATTTGATCTAATAACATAAAGAAAGATATGTA 189
QY 157 GAAAGAGCAATATACGATTTCTGCCATTTTACCGAGATATGTAAACCGAGGATGAG 216
DB 157 GAAAGAGCAATATACGATTTCTGCCATTTTACCGAGATATGTAAACCGAGGATGAG 216
QY 190 GTTATACAGGTAATGATTCGGGAGTAGCGCATGAAGAAATAGTACGACATCAATTG 249
DB 190 GTTATACAGGTAATGATTCGGGAGTAGCGCATGAAGAAATAGTACGACATCAATTG 249
QY 217 ATCAATGCAAGGAATGCAATTCCTGATTCATGAGAAATCCAGCGAGCAAAAGATTA 276
DB 217 ATCAATGCAAGGAATGCAATTCCTGATTCATGAGAAATCCAGCGAGCAAAAGATTA 276
QY 250 AAATAGATGTTCTAGATTAATGATTTTATAGGACAGATTTATTTACATGTTATG 309
DB 250 AAATAGATGTTCTAGATTAATGATTTTATAGGACAGATTTATTTACATGTTATG 309
QY 277 GGTTCGGAAGTGAATCGCTATACGACTTTTATGGTCAATTTATCCAAAATTTACGAGT 336
DB 277 GGTTCGGAAGTGAATCGCTATACGACTTTTATGGTCAATTTATCCAAAATTTACGAGT 336
QY 310 GTAGCTGTAACTGGTGCACATGTTAAACCTTCTACACAGCTTTATTTACATGTTATG 369
DB 310 GTAGCTGTAACTGGTGCACATGTTAAACCTTCTACACAGCTTTATTTACATGTTATG 369
QY 337 ATCGCTGTAACTGGTGCACATGTTAAACCTTCTACACAGCTTTATTTACATGTTATG 396
DB 337 ATCGCTGTAACTGGTGCACATGTTAAACCTTCTACACAGCTTTATTTACATGTTATG 396
QY 370 AATGGTGATAAAGACTTCAATTTTAAATGGTGATGGCAAGGATGGGATTCGCTGAA 429
DB 370 AATGGTGATAAAGACTTCAATTTTAAATGGTGATGGCAAGGATGGGATTCGCTGAA 429
QY 397 AGTGGTGTCGCTCCAAAGATTATCTAATTTGGAGATGGAACAGACATGTTGATCCGAA 456
DB 397 AGTGGTGTCGCTCCAAAGATTATCTAATTTGGAGATGGAACAGACATGTTGATCCGAA 456
QY 430 AGTGATTAATTCGCTTTTACGAGCATGTAATAGAGCTCACATTTTAAAGTTAAACCT 489
DB 430 AGTGATTAATTCGCTTTTACGAGCATGTAATAGAGCTCACATTTTAAAGTTAAACCT 489
QY 457 GCGGAATTCCTTTCAATTTGAGGCTGTGAATATCGCCGTCATTTCTTGCCTTATTCACCA 516
DB 457 GCGGAATTCCTTTCAATTTGAGGCTGTGAATATCGCCGTCATTTCTTGCCTTATTCACCA 516
QY 490 GATTACGCAATTAATGACAAATTTGATTTTCGATCATCTGATTTTAAAGATATTAAT 549
DB 490 GATTACGCAATTAATGACAAATTTGATTTTCGATCATCTGATTTTAAAGATATTAAT 549
QY 517 GACTATGCAATCATGACCAATTCGATTTTATGATCCGGAATTTATTAAGATATCGAC 576
DB 517 GACTATGCAATCATGACCAATTCGATTTTATGATCCGGAATTTATTAAGATATCGAC 576
QY 550 GATGTTTTGATGCAATTCGCAAAATGGCACAATAATGTTTAAAAAGGATTTATTTGCTGG 609
DB 550 GATGTTTTGATGCAATTCGCAAAATGGCACAATAATGTTTAAAAAGGATTTATTTGCTGG 609
QY 577 GATGTTTTGATGCAATTCGCAAAATGGCACAATAATGTTTAAAAAGGATTTATTTGCTGG 636
DB 577 GATGTTTTGATGCAATTCGCAAAATGGCACAATAATGTTTAAAAAGGATTTATTTGCTGG 636
QY 610 GGTGATGATGACATCTAGTAAATTTGAAGCAGATGTTCCAAATTTATTTATGATTT 669
DB 610 GGTGATGATGACATCTAGTAAATTTGAAGCAGATGTTCCAAATTTATTTATGATTT 669
QY 637 GGTGATGATGATTTATTTACGCAATTTAAAGCAAAATGTCGCGATTTATTTATGCTGTC 696
DB 637 GGTGATGATGATTTATTTACGCAATTTAAAGCAAAATGTCGCGATTTATTTATGCTGTC 696
QY 670 AAAGATTCGATGATTTATGCTCAAAATTTCAATTTACGATTAAGGATTAAGGATTT 729
DB 670 AAAGATTCGATGATTTATGCTCAAAATTTCAATTTACGATTAAGGATTAAGGATTT 729
QY 697 ACAGAAACGATGATATCCAGACGAAACATCGAACGCAACACTTCTGCTTCTGCTTTT 756
DB 697 ACAGAAACGATGATATCCAGACGAAACATCGAACGCAACACTTCTGCTTCTGCTTTT 756

```

```

QY 730 GATGCTATGTTGGATGATGATTTTATGATGATCACTTCTCTCCCAATATGTTGATCAT 789
DB 730 GATGCTATGTTGGATGATGATTTTATGATGATCACTTCTCTCCCAATATGTTGATCAT 789
QY 757 GATGCTATGATGATGATGATTTTGGGCACTTCTCTGATGATGATGATGATGATGATGAT 816
DB 757 GATGCTATGATGATGATGATTTTGGGCACTTCTCTGATGATGATGATGATGATGATGAT 816
QY 790 ACAGTTTAAATGATGATGATGATTTTAAATGATGATGATGATGATGATGATGATGAT 845
DB 790 ACAGTTTAAATGATGATGATGATTTTAAATGATGATGATGATGATGATGATGATGAT 845
QY 817 GATATCTTAAATGATGATGATTTTAAATGATGATGATGATGATGATGATGATGATGAT 872
DB 817 GATATCTTAAATGATGATGATTTTAAATGATGATGATGATGATGATGATGATGATGAT 872

RESULT 15
US-08-940-572-3/c
; Sequence 3, Application US/08940572
; Patent No. 6310193
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; APPLICANT: Black, Michael T.
; APPLICANT: Hodgson, John E.
; APPLICANT: Knowles, David J.
; APPLICANT: Lonetto, Michael A.
; APPLICANT: Nicholas, Richard O.
; APPLICANT: Stodola, Robert K.
; TITLE OF INVENTION: No. 6310193el MurC
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/940,572
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024022
; FILING DATE: 16-AUG-1996
; APPLICATION NUMBER: US 08/889711
; FILING DATE: 08-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Q. Todd
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50533-04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215/994-2252
; TELEFAX: 215/994-2222
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-08-940-572-3

Query Match      23.6%; Score 310.4; DB 4; Length 1825;
Best Local Similarity 59.0%; Pred. No. 5e-63;
Matches 607; Conservative 0; Mismatches 411; Indels 10; Gaps 4;

QY 264 AAGTTAATGATTTTATGACACAGATTTATTTGATCAATATATCTTCACTAGCTGTAAC 323
DB 264 AAGTTAATGATTTTATGACACAGATTTATTTGATCAATATATCTTCACTAGCTGTAAC 323
QY 1797 ACCTTACCATGAGTTTCTAGTAGCTTTATTCGCTGACTTTTGTAGTACGAGTAGCAGG 1738
DB 1797 ACCTTACCATGAGTTTCTAGTAGCTTTATTCGCTGACTTTTGTAGTACGAGTAGCAGG 1738
QY 324 TGCAATGGAAGTCTTCAACAGGTTTATTTATCACTGTTTATGATGATGATGATGATGAT 383
DB 324 TGCAATGGAAGTCTTCAACAGGTTTATTTATCACTGTTTATGATGATGATGATGATGAT 383
QY 1737 AGCAGATGGAAGTCTTCAACAGGTTTATTTGCTCTCATGCTTCTCTCATGATGATGAT 1678
DB 1737 AGCAGATGGAAGTCTTCAACAGGTTTATTTGCTCTCATGCTTCTCTCATGATGATGAT 1678
QY 384 GACTTCAATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 443
DB 384 GACTTCAATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 443

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2004, 02:47:10 ; Search time 5352 Seconds
(without alignments)
10673.789 Million cell updates/sec

Title: US-10-712-713-1
Perfect score: 1318
Sequence: 1 atgacacacatacttttgt.....gaaaatgcgtttaagctt 1318

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 50

Total number of hits satisfying chosen parameters: 17

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1314	99.7	1335	6	AX742069 Sequence
2	1314	99.7	2424	6	AR354274 Sequence
3	1311	99.5	1332	6	AX617936 Sequence
4	1263	95.8	1335	6	AX742071 Sequence
5	1212	92.0	301550	1	AP003134 Sequence
6	1212	92.0	342600	1	AP003134 Staphyloc
7	1161	88.1	272850	1	AP003134 Staphyloc
8	995	75.5	1314	1	AF034076 Staphyloc
9	995	75.5	1314	6	AX191741 Sequence
10	842	63.9	1351	6	E35613 MurC. 6/200
11	842	63.9	1351	6	BD178030 MurC. 4/2
12	434	32.9	660	6	E35614 MurC. 6/200
13	434	32.9	660	6	BD178031 MurC. 4/2
14	275	20.9	619	6	AR194630 Sequence
15	90	6.8	888	6	AR149358 Sequence
16	90	6.8	888	6	AR405068 Sequence
17	90	6.8	888	6	AR411614 Sequence

ALIGNMENTS

RESULT 1	AX742069	Sequence 27 from Patent WO03025007.	1335 bp	DNA	linear	PAT 10-MAY-2003
LOCUS	AX742069	Sequence 27 from Patent WO03025007.				
DEFINITION	AX742069	Sequence 27 from Patent WO03025007.				
ACCESSION	AX742069	Sequence 27 from Patent WO03025007.				
VERSION	AX742069.1	GI:30524566				
KEYWORDS	Staphylococcus aureus					
SOURCE	Staphylococcus aureus					
ORGANISM	Staphylococcus aureus					
REFERENCE	1	Bacteria; Firmicutes; Bacillales; Staphylococcus.				
AUTHORS	Edwards, A., Dharamsi, A., Vedadi, M., Alam, M.Z., Awrey, D., Beattie, B., Domagala, M., Houston, S., Kanagarajah, D., Nethery, K., NG, I., Mansoury, K., McDonald, M.L., Pinder, B., Viola, C. and Wrezel, O.					
TITLE	Novel purified polypeptides involved in membrane biosynthesis					
JOURNAL	Patent: WO 03025007-A. 27 27-MAR-2003,					
FEATURES	Affinium Pharmaceuticals, Inc. (CA)					
source	Location/Qualifiers					
	1. 1335					
	/organism="Staphylococcus aureus"					
	/mol_type="genomic DNA"					
	/db_xref="taxon:1280"					
ORIGIN						
Query Match	99.7%	Score 1314;	DB 6;	Length 1335;		
Best Local Similarity	100.0%;	Pred. No. 0;				
Matches 1314;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
QY	1	ATGACACACTATCATTTTGTTCGGAAATTAAGGTTCTGGCATGAGTTTACGCAAAATC	60			
DB	22	ATGACACACTATCATTTTGTTCGGAAATTAAGGTTCTGGCATGAGTTTACGCAAAATC	81			
QY	61	ATGCATGATTTAGGACATGAAGTTTCAGGATTCGATATTCGACTAGTATTACAGAA	120			
DB	82	ATGCATGATTTAGGACATGAAGTTTCAGGATTCGATATTCGACTAGTATTACAGAA	141			
QY	121	GTTCCTTTAGAAATAAGGGGATAAAAATATTACCTTTGATGCTAATAACATAAAGAA	180			
DB	142	GTTCCTTTAGAAATAAGGGGATAAAAATATTACCTTTGATGCTAATAACATAAAGAA	201			
QY	181	GATATCGTAGTTATACAAAGTAATGCTTCGAGTAGCCATGAAGAAATAGTAGTGCA	240			

Db 202 GATATGGTAGTTATACAGGTAATGCAATTCGGAGTAGCCATGAAGAATAAGTAGCTGCA 261
 QY 241 CATCAATTGAAATAGATGTTGTAAGTTATATAGATTTTATAGGACAGATTAATGATCAA 300
 Db 262 CATCAATTGAAATAGATGTTGTAAGTTATATAGATTTTATAGGACAGATTAATGATCAA 321
 QY 301 TATACCTCAGTAGCTGTAACCTGTCACATGTAAGTAACTTCTACACAGGTTTATATCA 360
 Db 322 TATACCTCAGTAGCTGTAACCTGTCACATGTAAGTAACTTCTACACAGGTTTATATCA 381
 QY 361 CATGTTATGAATGGTGATATAAAGAACTTCATTTTAAATTTGGTGATGGCACAGGTATGGA 420
 Db 382 CATGTTATGAATGGTGATATAAAGAACTTCATTTTAAATTTGGTGATGGCACAGGTATGGA 441
 QY 421 TTGCTGAAAGTGAATTTTTCGCTTTTCAGGATGTAATAGACGTACCTTTTAAAGT 480
 Db 442 TTGCTGAAAGTGAATTTTTCGCTTTTCAGGATGTAATAGACGTACCTTTTAAAGT 501
 QY 481 TATAAACCCTGATACGCAATTTATGACAAATATTGATTTTCGATCATCTCTGATTTTAAA 540
 Db 502 TATAAACCCTGATACGCAATTTATGACAAATATTGATTTTCGATCATCTCTGATTTTAAA 561
 QY 541 GATATTAATGATGTTTTCGATCATCTCAAGAAATGGCACATAATGTTTAAAAAGGTATT 600
 Db 562 GATATTAATGATGTTTTCGATCATCTCAAGAAATGGCACATAATGTTTAAAAAGGTATT 621
 QY 601 ATTGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 Db 622 ATTGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 681
 QY 661 TATGGATTTAAAGATTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 Db 682 TATGGATTTAAAGATTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 741
 QY 721 ACTGCTTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 Db 742 ACTGCTTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 801
 QY 781 GGTGACCATACAGTTTAAATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 Db 802 GGTGACCATACAGTTTAAATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 861
 QY 841 GATGTTTACAAATTTAAAGAAAGCAATGAAACGTTTGGTGGTGTAAACGTCGTTTCAAT 900
 Db 862 GATGTTTACAAATTTAAAGAAAGCAATGAAACGTTTGGTGGTGTAAACGTCGTTTCAAT 921
 QY 901 GAACTACCAATTCGAATCAAGTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 Db 922 GAACTACCAATTCGAATCAAGTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 981
 QY 961 AGTGCTACCAATTTGAAACAGCACGAAAGAAATATCCACATAAAGAAAGTTGTTGCAAGTATT 1020
 Db 982 AGTGCTACCAATTTGAAACAGCACGAAAGAAATATCCACATAAAGAAAGTTGTTGCAAGTATT 1041
 QY 1021 CAACCAACACATTTCTCTAGAACACAGGCAATTTTAAATGAAATTTGCAAGAAAGTTTAAAGT 1080
 Db 1042 CAACCAACACATTTCTCTAGAACACAGGCAATTTTAAATGAAATTTGCAAGAAAGTTTAAAGT 1101
 QY 1081 AAAGCAGATCGTGATTTCTCTAGAACACAGGCAATTTTAAATGAAATTTGCAAGAAAGTTTAAAGT 1140
 Db 1102 AAAGCAGATCGTGATTTCTCTAGAACACAGGCAATTTTAAATGAAATTTGCAAGAAAGTTTAAAGT 1161
 QY 1141 TTAACGATACAGATTTAAATTTGATTAATTTGAAAGGTGATCGTTTAAATTTAAATGAAAGTCT 1200
 Db 1162 TTAACGATACAGATTTAAATTTGATTAATTTGAAAGGTGATCGTTTAAATTTAAATGAAAGTCT 1221
 QY 1201 ATTATGATTTAGAACAAATTTGATTAATTTGATTAATTTGATTAATTTGATTAATTTGATTAAT 1260
 Db 1222 ATTATGATTTAGAACAAATTTGATTAATTTGATTAATTTGATTAATTTGATTAATTTGATTAAT 1281
 QY 1261 CAAAAATTTACAAATGCAATTTAGATAAATTTAGGCAATGAAATGCGTTTAA 1314
 Db 1282 CAAAAATTTACAAATGCAATTTAGATAAATTTAGGCAATGAAATGCGTTTAA 1335

RESULT 2

AR354274

LOCUS

DEFINITION

AR354274

ACCESSION

AR354274

VERSION

AR354274.1

KEYWORDS

Unknown.

ORGANISM

Unclassified.

REFERENCE

1 (bases 1 to 2424)

AUTHORS

Rosen, C.A.

TITLE

Staphylococcus aureus polynucleotides and sequences

JOURNAL

Patent: US 6593114-A 392 15-JUL-2003;

FEATURES

Location/Qualifiers

1..2424

source

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match

Best Local Similarity 99.7%; Score 1314; DB 6; Length 2424;

Matches 1314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACACTATCATTTTCTCGGAATTAAGGTTCTGGCATGAGTTTCATTAGCACAAATC 60

Db 627 ATGACACACTATCATTTTCTCGGAATTAAGGTTCTGGCATGAGTTTCATTAGCACAAATC 686

QY 61 ATGCATGATTTAGGACATGAACTTCAAGCATCGGATATTGGAACCTACGCTATTTCAGAA 120

Db 687 ATGCATGATTTAGGACATGAACTTCAAGCATCGGATATTGGAACCTACGCTATTTCAGAA 746

QY 121 GTTGCTCTTAGAAATTAAGGGGATAAAAAATATACCATTTTGATGCTTAATAACATAAAGAA 180

Db 747 GTTGCTCTTAGAAATTAAGGGGATAAAAAATATACCATTTTGATGCTTAATAACATAAAGAA 806

QY 181 GATATGGTAGTTATACAGGTAATGCAATTCGGAGTAGCCATGAAGAAATAGTACGTGCA 240

Db 807 GATATGGTAGTTATACAGGTAATGCAATTCGGAGTAGCCATGAAGAAATAGTACGTGCA 866

QY 241 CATCAATTTGAAATAGATGTTGTAAGTTATATATGATTTTATAGGACAGATTAATGATCAA 300

Db 867 CATCAATTTGAAATAGATGTTGTAAGTTATATATGATTTTATAGGACAGATTAATGATCAA 926

QY 301 TATACCTCAGTAGCTGTAACCTGGTGCAATGTTGTAAGAACTTCTACACAGGTTTATTTATCA 360

Db 927 TATACCTCAGTAGCTGTAACCTGGTGCAATGTTGTAAGAACTTCTACACAGGTTTATTTATCA 986

QY 361 CATGTTATGAATGGTGATATAAAGAACTTCATTTTAAATTTGTTGATGGCACAGGTATGGA 420

Db 987 CATGTTATGAATGGTGATATAAAGAACTTCATTTTAAATTTGTTGATGGCACAGGTATGGA 1046

QY 421 TTGCTGAAAGTGAATTTTTCGCTTTTCAGGCAATGTAATAGACGTCACTTTTAAAGT 480

Db 1047 TTGCTGAAAGTGAATTTTTCGCTTTTCAGGCAATGTAATAGACGTCACTTTTAAAGT 1106

QY 481 TATAAACCCTGATACGCAATTTATGACAAATATTGATTTTCGATCATCTCTGATTTTAAA 540

Db 1107 TATAAACCCTGATACGCAATTTATGACAAATATTGATTTTCGATCATCTCTGATTTTAAA 1166

QY 541 GATATTAATGATGTTTTCGATGCAATTCAGAAATGGCACATAATGTTTAAAAAGGTATT 600

Db 1167 GATATTAATGATGTTTTCGATGCAATTCAGAAATGGCACATAATGTTTAAAAAGGTATT 1226

QY 601 ATTGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660

Db 1227 ATTGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1286

QY 661 TATGATTTTAAAGATTCGGATGACATTTATGCTCAAAATATTTCAAATTTACGATTAAGGT 720

Db 720 TATGATTTTAAAGATTCGGATGACATTTATGCTCAAAATATTTCAAATTTACGATTAAGGT 720

Db 1162 TTAACGATACAGATTAAATGATAAATGAAGTGCATCGTTAAATGAAGATTCT 1221
QY 1201 ATTAAGTATTAGAACAAATTTGATATGCTGTTATTTATTTATGCGTGCAGGTGATATT 1260
Db 1222 ATTAATGATTAGAACAAATTTGATATGCTGTTATTTATTTATGCGTGCAGGTGATATT 1281
QY 1261 CAAAATTTACAAAATGCAATTTAGATAAATAGGCATGAAAATCGGTTT 1311
Db 1282 CAAAATTTACAAAATGCAATTTAGATAAATAGGCATGAAAATCGGTTT 1332

RESULT 4
AX742071
LOCUS AX742071 1335 bp DNA linear PAT 10-MAY-2003
DEFINITION Sequence 29 from Patent WO03025007.
ACCESSION AX742071
VERSION AX742071.1 GI:30524567
KEYWORDS
SOURCE Staphylococcus aureus
ORGANISM Staphylococcus aureus
REFERENCE 1
AUTHORS Edwards, A., Dharmsi, A., Vedrali, M., Alam, M.Z., Awrey, D., Beattie, B., Domagala, M., Houston, S., Kanagarajah, D., Nethery, K., Ng, I., Mansoury, K., McDonald, M.L., Pinder, B., Viola, C. and Wrezel, O.
TITLE Novel purified polypeptides involved in membrane biosynthesis
JOURNAL Patent: WO 03025007-A 29 27-MAR-2003;
Affinium Pharmaceuticals, Inc. (CA)
FEATURES
source Location/Qualifiers
1..1335
/organism="Staphylococcus aureus"
/mol_type="genomic DNA"
/db_xref="taxon:1280"

ORIGIN

Query Match 95.8%; Score 1263; DB 6; Length 1335;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1313; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACACACTATCATTTTGTGCGGAATTAAGGTTCTGCGCATGATTCATAGCACAAATC 60
Db 22 ATGACACACTATCATTTTGTGCGGAATTAAGGTTCTGCGCATGATTCATAGCACAAATC 81
QY 61 ATGCATCATTTAGACATGAAGTTCAAGGATCGGATATTGAGAACTAGTATTTACAGAA 120
Db 82 ATGCATCATTTAGACATGAAGTTCAAGGATCGGATATTGAGAACTAGTATTTACAGAA 141
QY 121 GTTGCTCTTGAATAAGGGGATAAAAATATTACCATTTGATGCTTAATAACATAAAAGAA 180
Db 142 GTTGCTCTTGAATAAGGGGATAAAAATATTACCATTTGATGCTTAATAACATAAAAGAA 201
QY 181 GATATGCTAGTTATACAGGATTAATGCTTCGCGATAGCCATGAAGAAATAGTACGTGCA 240
Db 202 GATATGCTAGTTATACAGGATTAATGCTTCGCGATAGCCATGAAGAAATAGTACGTGCA 261
QY 241 CATCAATTTGAAATAGATGTTGTAAGTTATAATGCAATTTTATAGGACAGATTTATGATCAA 300
Db 262 CATCAATTTGAAATAGATGTTGTAAGTTATAATGCAATTTTATAGGACAGATTTATGATCAA 321
QY 301 TATACCTCAGTAGCTGTAACCTGCTGCAATGGTAAACCTTCTACAAAGGTTTATTTATCA 360
Db 322 TATACCTCAGTAGCTGTAACCTGCTGCAATGGTAAACCTTCTACAAAGGTTTATTTATCA 381
QY 361 CATGTTATGATGCTGTAATAAGACTTCAATTTTAAATGCTGATGCGACAGGTATGGGA 420
Db 382 CATGTTATGATGCTGTAATAAGACTTCAATTTTAAATGCTGATGCGACAGGTATGGGA 441
QY 421 TTGCTTGAAGTGAATTTTTCGCTTTTGGGCAATGTAATATAGACGTCACCTTTTAAAT 480
Db 442 TTGCTTGAAGTGAATTTTTCGCTTTTGGGCAATGTAATATAGACGTCACCTTTTAAAT 501

QY 481 TATAAACCCTGATTACGCAATTTATGCAAAATATTGATTTCCGATCATCTGATTATTTAA 540
Db 502 TATAAACCCTGATTACGCAATTTATGCAAAATATTGATTTCCGATCATCTGATTATTTAA 561
QY 541 GATATTAATGATGTTTGTGATGCAATCCAGAAATGGCACAATAATGTTTAAAAAGGTTAT 600
Db 562 GATATTAATGATGTTTGTGATGCAATCCAGAAATGGCACAATAATGTTTAAAAAGGTTAT 621
QY 601 ATTGCTTGGGTTGATGATGAACATCTACGTAAATTTGACAGATGTTTCCCAATTTATAT 660
Db 622 ATTGCTTGGGTTGATGATGAACATCTACGTAAATTTGACAGATGTTTCCCAATTTATAT 681
QY 661 TATGGAATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAAATACGATAAAGGT 720
Db 682 TATGGAATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAAATACGATAAAGGT 741
QY 721 ACTGCTTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 780
Db 742 ACTGCTTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 801
QY 781 GGTGACCATACAGTTTAAATGCAATGCTGTTAAATGCGATGTTAGTATTTAGAGAAGCTA 840
Db 802 GGTGACCATACAGTTTAAATGCAATGCTGTTAAATGCGATGTTAGTATTTAGAGAAGCTA 861
QY 841 GATGTTTACAAATTTAAAGAAAGCAATTTAGAAACGTTTGGTGTGTTAAACGTCGTTTCAAT 900
Db 862 GATGTTTACAAATTTAAAGAAAGCAATTTAGAAACGTTTGGTGTGTTAAACGTCGTTTCAAT 921
QY 901 GAAACTACAAATTTGCAAAATCAAAGTTTATGTTAGATGATTTATGATGATGATTTATGATGAT 960
Db 922 GAAACTACAAATTTGCAAAATCAAAGTTTATGTTAGATGATTTATGATGATGATTTATGATGAT 981
QY 961 AGTGCTTACAAATTTGAAACAGCAGCAAGAAATATCCACATAGAAAGTTGTTGCAATTT 1020
Db 982 AGTGCTTACAAATTTGAAACAGCAGCAAGAAATATCCACATAGAAAGTTGTTGCAATTT 1041
QY 1021 CAACCAACACACTTTCTCTAGAACACAGGCAATTTTAAATGCAATTTGCAAGAAAGTTTAAAT 1080
Db 1042 CAACCAACACACTTTCTCTAGAACACAGGCAATTTTAAATGCAATTTGCAAGAAAGTTTAAAT 1101
QY 1081 AAAGCAGATCGTGTATTTCTTATGTTGAAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1140
Db 1102 AAAGCAGATCGTGTATTTCTTATGTTGAAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1161
QY 1141 TTAAGATACAGATTTAATGATAAATTTGAAGTGCATCGTTAAATTAATGAAGATCT 1200
Db 1162 TTAAGATACAGATTTAATGATAAATTTGAAGTGCATCGTTAAATTAATGAAGATCT 1221
QY 1201 ATTAATGATTAGAACAAATTTGATATGCTGTTATTTATTTATGCGTGCAGGTGATATT 1260
Db 1222 ATTAATGATTAGAACAAATTTGATATGCTGTTATTTATTTATTTATGCGTGCAGGTGATATT 1281
QY 1261 CAAAATTTACAAATGCAATTTAGATAAATAGGCATGAAAATCGGTTTAA 1314
Db 1282 CAAAATTTACAAATGCAATTTAGATAAATAGGCATGAAAATCGGTTTAA 1335

RESULT 5
AP003134/c
LOCUS AP003134 301550 bp DNA linear BCT 24-APR-2003
DEFINITION Staphylococcus aureus subsp. aureus N315 genomic DNA, complete genome, section 6/10.
ACCESSION AP003134 BA000018
VERSION AP003134.2 GI:14349226
KEYWORDS
SOURCE Staphylococcus aureus subsp. aureus N315
ORGANISM Staphylococcus aureus subsp. aureus N315
REFERENCE 1
AUTHORS Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M., Matsumaru, H., Maruyama, A., Murakami, H., Hosoyama, A., Mizutani, U., Takahashi, N.K., Sawano, T., Inoue, R., Kaito, C.,

AUTHORS

Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Iian, J., Ito, T., Kanamori, M., Matsumaru, H., Maruyama, A., Murakami, H., Hosoyama, A., Mizutani, U. Y., Takahashi, N. K., Sawano, T., Inoue, R., Kaito, C., Sekimizu, K., Hirakawa, H., Kuhara, S., Goto, S., Yabuzaki, J., Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C., Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. and Hiramatsu, K. Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*

TITLE

JOURNAL Lancet 357 (9264), 1225-1240 (2001)

MEDLINE

PUBMED 21311952

REFERENCE

2 (bases 1 to 342600)

AUTHORS

Ohta, T.

TITLE

Submitted (28-FEB-2001) Toshiko Ohta, University of Tsukuba College of Medical Technology and Nursing, Department of Medical Technology; 1-1-1 Ten-nodai, Tsukuba, Ibaraki 305-8577, Japan (E-mail: tohta@akura.cc.tsukuba.ac.jp, Tel: 81-298-53-3454, Fax: 81-298-53-3454)

On May 29, 2001 this sequence version replaced gi:13875943.

COMMENT

Location/Qualifiers

source

1. .342600

/organism="Staphylococcus aureus subsp. aureus Mu50"

/mol_type="genomic DNA"

/strain="Mu50"

/sub_species="aureus Mu50"

/db_xref="taxon:158878"

complement (42. .815)

/gene="SAV1628"

complement (42. .815)

/gene="SAV1628"

/codon_start=1

/transl_table=11

/product="conserved hypothetical protein"

/protein_id="BAB57790.1"

/db_xref="GI:14247400"

/translation="MKHGFNRSLALGQGLDLKQQTIVVILGVGVGSFAEALART NIGHILDKDDVDITNQRHLATTTVQSGKSLMEIRIKLINPDKCVPLMFYET EYEEIFDNDIDYFDASDTIMYKVLMEKCLDRGIVISMGAAKNTDPTRTIS DISKTHDPMAKIIRQKLKLGIRKGIPIVVFSDPSPIREDVKIVGDKNALNRKGQ MPSSNAFVPSVVGICASYVNDVLAIDIPVRIRKDGQ"

complement (1029. .1130)

/gene="SAV1629"

complement (1029. .1130)

/gene="SAV1629"

/codon_start=1

/transl_table=11

/product="hypothetical protein"

/protein_id="BAB57791.1"

/db_xref="GI:14247401"

/translation="MGFLAAVTPYMKDVQNKQKQTHLYTGRMIKLL"

complement (1276. .3042)

/gene="asps"

complement (1276. .3042)

/gene="asps"

/note="SAV1630"

/codon_start=1

/transl_table=11

/product="aspartyl-tRNA synthetase"

/protein_id="BAB57792.1"

/db_xref="GI:14247402"

/translation="MSKRTTYCVGLVTRAFILGQELTLKGWNNRDLGLIFVLDLRE GIVQVNPFAFSEBALAIETVSEYVVEVQGTWKDPTVNPVKIKTGQVEQVTNI KVINKSTPTPFSNEENNVNDENIRLRYLDLRQELAQTFKMRHQITRSIRQLDD EGFDLPTVLTSTPGARDYLPVRHDGEFYALPQSPQLKMLISGFDKYQI KFCRDLRADRQPEFTQDIEMSFVDQSDVMQGEMLKKVVKVKGVEINGAFPR MYTKEARRYSKDPDRFQWELIDVSLGDMDFKFKOTVENDGHIKAVKGAEE QYTRKMDALTEFVNIYKAGLAWKVVEGLTGPIGRFETENVELLITUTGAEAG LVMFVADKPNVAQSLGALRVKLAKELIDETKLNFLVTDWPLLEDYEDDAKRYAA HHFTSPKADIAKGTAPAEAEANAYDIVLNGYELGGGIRIHGDELQEQMFVLGF TREQAQEQFGLDADFYGAPPHPGGIALGLDLRLVMLLTNRNLRTDTAFPKTASATCL LTNAPGEVSKQLEELSLIRH"

gene

CDS

complement (3058. .4320)

/gene="hiss"

complement (3058. .4320)

/gene="hiss"

/note="SAV1631"

/codon_start=1

/transl_table=11

/product="histidyl-tRNA synthetase"

/protein_id="BAB57793.1"

/db_xref="GI:14247403"

/translation="MKIPRGTDILPEDSKKRYIENQDLMTFYNYKEIRTIPE STDLPARGVGDSTDVQKEMTYTFKQDRSITLRPEGTAAVVRSYIEHKMGQNPQPI KLYNGPMFYERKQGRYQFQVRAIGAENPSVDAVLAWVHIYQSGFKHLK LIVINSVDMASRKEYNELVHFEFVHIEFSCDCQSLHNPWRILDCVDRKEALK TAPRTDPLNEESKAYVQKAYLDDGLGPIYIEDNLVRLGDTTITAFELMDNPNY DGAITLCCGRYNGLLSDGSETGIGFALSIERLLALBEEGLIEENLDIFI VMKIDVNMVTGSETIBLDALVEYFK"

complement (4782. .5657)

/gene="lyth"

complement (4782. .5657)

/gene="lyth"

/note="SAV1632"

/codon_start=1

/transl_table=11

/product="N-acetylmuramoyl-L-alanine amidase"

/protein_id="BAB57794.1"

/db_xref="GI:14247404"

/translation="MKKIEAWLSKGLKNKRTLIIVIAFVLPFIIFLLNSDPSG NITITENAEALRTGPAAPVIVYKVEKGDFKKIKGVKMWIEVEDTSSNEKGTAGWHT NLDIVADNTKEKNPLOGKTI VLDPCGGSDGASNTKYKSLBKDYTLKAKELQRTL EKSGATVMTFTDITYSLENDIKGDALYSIHNDALSSNANGMTVVIYHNDQALAL DTLDATQKGLLSNRGRQENYQVLRQTKVPAVLLELGYISNPTDTMTIKQLHRQI LRQAIVDGLKIYFSA"

complement (5654. .6106)

/gene="SAV1633"

complement (5654. .6106)

/gene="SAV1633"

/codon_start=1

/transl_table=11

/product="conserved hypothetical protein"

/protein_id="BAB57795.1"

/db_xref="GI:14247405"

/translation="MKVVQVRVKEASVTNLTNNQIKKGYCLLVIGIGNSTEQDADVI AKIANARLPEDNNKLNFIQMNQGEILSVQFTLYADVKKRNPGRFSGNSKNPDQAV KIYEPNDALRAYGLTVKTGEFGTHMNVISINNDGFPVTIYESQDKIQ"

complement (6118. .8307)

/gene="rela"

complement (6118. .8307)

/gene="rela"

/note="SAV1634"

/codon_start=1

/transl_table=11

/product="GTP pyrophosphokinase"

/protein_id="BAB57796.1"

/db_xref="GI:14247406"

/translation="MNNPEYPSADEVLHKAQSYLSADEYEVYLVKSYHAYEAHKQGRF KNGLPYIMHPIQVAGILTEMRLDPTIVAGFLHVDITPTFEDVKEMNEEKVARIV DGVTLKKVYRSKEEQAEHRKLFIAIAKQVRVILVKLADRLHNMETLKAMPREKQ IIRSETLEIYAPLAHLGINTIKWELEDTALAYIDNVQYFRIVNLKMKCSREAYI ETADRIRTEMDRMINEGDIRPHIYIIRKMMKQKQFQDQIFDLALRVINSN DCIALGLVHTLTMKPGPFQDIYAMPKQNLQSLTFTVVGPDGLFQIDLLKVDLOS TAEVAAHAWAYKEGKYSKQDTYQNKLNWLKELAEADHTSDAQEFMELTKYDQS DKVYFTPASDVIELPYGAVPDIYFAIHSVEGNKMIKAKVNGKI VPIDYLTQGDIV EIRTSKHSVSPSRDLKTVKSSAKGIKSPFKQDRSNIIEKGRMVEVEIKQGRPR VEDILTEKNIQVNEKYFANEDDLPAAVGVGVTLSQIVNKLTEQRILDKQALNE EQBVTSLPDKNIIITDSGVLEENVLISKCCNPFGDDIVGYITKGHGIKVHR TDCPNKINTERLINISVWVKNDVIRVVEKIKQLGDVYVTVTVRN"

complement (8735. .9253)

/gene="apt"

complement (8735. .9253)

/gene="apt"

```
/note="SAV1635"
/codon_start=1
/transl_table=11
/product="adenine phosphoribosyl transferase"
/protein_id="BAB57797.1"
/db_xref="GI:14247407"
/translation="MDLKQYSEVODWPKPGVSFKDITTIMDNGEAYGYATDKLVEYA
KRDVDIVVGPARGFIIICPVAYSNGIGFAPVRKEGKLPREVIRIYEDLEYGTNVLIT
MKHDAIKGQVRVLIITDDLLATGGTIEAIAIKLVEKUGGIIVVGAFFIELKYLINGIEKIK
DIDVMSLISYDE"
gene complement(9275..11548)
/feature="SAV1636"
CDS complement(9275..11548)
/gene="SAV1636"
/note="similar to single-strand DNA-specific exonuclease"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57798.1"
/db_xref="GI:14247408"
/translation="MIKPKYKWKLTAKPAEYISDELTSKLKLPVVKKILLESKSIIDDEQ
AIESIISDTHNDALQISDMTKAIEKIKRAIANDEKILVIGDYDADGVSTTILVT
LQLGAQYGHIPNFTFGYGNELAFNAHDEGILITVDNGIQHNEIKWODLG
VDVITDHEIGSTLPEAVAI VHPMHPSPNVPQOLCGAGVAYKLAQALIENVPDYFK
ALVAGTTADIVSLTDENRSIVKQSLKVLNDQCPSTVKALKKAGNDINIBETIGFI
IGRLNAYGRLLDASLACELMLDDEERAAFLAEQVHFNRRKDIIVATITBEANAMA
ETKVKGDLFLLAKENHGVGLIVASKIVETFPALPTILINIDREONHAKGSARSID
QVSFEILSAQHEITAKGEGHMAAGMTMDIENIRSLAELNKKMKLSKTSLDPVK
FVLLTENDITIKIRDMNRLRPFPTDSRIFEMDDLSVSKVAIGQKNHKLKLT
GESINIAALFPWNGHLEPLODEOPINILGSVOINWNGNOSPQLIQLDAMNEQILID
YRSKSPSPFTENDENIVLILHPKSDKNVANYEYGEIEIKQTDKVLNLDLSTMEDL
SNSIQLOFQSLQYIVLQNHHSIYFDGIPNMDIIFKCYKALITKQETNITQKGMJLCOH
LSVAPDITKFMKVPDLDFKVFQEDGLIRINQPKRSDSSKVTQLRQORMDVEKQL
LYQDFSEIKNWKISQLS"
gene complement(11751..14030)
CDS complement(11751..14030)
/gene="secP"
/note="SAV1637"
/codon_start=1
/transl_table=11
/product="protein-export membrane protein SecDF"
/protein_id="BAB57799.1"
/db_xref="GI:14247409"
```

Query Match 92.0%; Score 1212; DB 1; Length 342600;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1312; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATGACACACTATCATTTTGTGCGAATTAAGGTTCTGCGCATGAGTTCAATAGCACAAATC	60
Db	132799	ATGACACACTATCATTTTGTGCGAATTAAGGTTCTGCGCATGAGTTCAATAGCACAAATC	132740
QY	61	ATGCATCATTTAGCACATGAGTTCAAGGATCGGATATGAGAACTAGTATTACAGAA	120
Db	132739	ATGCATCATTTAGCACATGAGTTCAAGGATCGGATATGAGAACTAGTATTACAGAA	132680
QY	121	GTTCCTCTTGAATAAGGGGATAAAAATATFACCATTTGATGCTTAATAACATAAAGAA	180
Db	132679	GTTCCTCTTGAATAAGGGGATAAAAATATFACCATTTGATGCTTAATAACATAAAGAA	132620
QY	181	GATATGGTAGTTATACAGGTAATATGCAITTCGCGAGTAGCCATGAAGAATAAGTAGTGCA	240
Db	132619	GATATGGTAGTTATACAGGTAATATGCAITTCGCGAGTAGCCATGAAGAATAAGTAGTGCA	132560
QY	241	CATCAATTGAATTAGATGTTGTAAGTTAATGATTTTTTTAGGACAGATTATTGATCAA	300
Db	132559	CATCAATTGAATTAGATGTTGTAAGTTAATGATTTTTTTAGGACAGATTATTGATCAA	132500
QY	301	TATACCTTCAGTAGCTGTAACCTGGTGCACATGTTAAACTTCTACAAAGGTTTATATCA	360
Db	132499	TATACCTTCAGTAGCTGTAACCTGGTGCACATGTTAAACTTCTACAAAGGTTTATATCA	132440

RESULT 7

AP004828/c

LOCUS

DEFINITION

AP004828

Staphylococcus aureus subsp. aureus MW2, section 7/10.

ACCESSION

AP004828.1

VERSION

AP004828.1

KEYWORDS

272850 bp DNA linear BCT 20-DEC-2002

Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,

strain:MW2, section 7/10.

AP004828 BA000033

AP004828.1 GI:21204850

AP004828.1

AP004828.1

AP004828.1

AP004828.1


```
/gene="MW1687"  
/note="ORFID:MW1687  
thioredoxin homolog"  
/codon_start=1  
/transl_table=11  
/protein_id="BAB95552.1"  
/db_xref="GI:21204857"  
/translation="MKQLSEQQFESLKQAGTVEFTAGWCPDRCRVIEPDLPELEARY  
PMFDFVSDRDRFMDICIEINGIPSFIVYKNGELIGSVIGKERKSIIEQIDAFIAY  
V"  
complement (9612..10688)  
/gene="MW1688"  
complement (9612..10688)  
/gene="MW1688"  
/note="ORFID:MW1688  
endo-1,4-beta-glucanase homolog"  
/codon_start=1  
/transl_table=11  
/protein_id="BAB95553.1"  
/db_xref="GI:21204858"  
/translation="MNINKVTLQRIOTLTLHGAPGFEEVKNYMTQMAPYVDEPI  
ENRMGGFFGVKSKNPNKRVMTAAHMDIEGPMITNITKNGMIQFTNLGVANDIMQG  
QRLVINKRNGDKLIGVSNIPKHFTGSGEAPHIKLTLDIGAQNEDEVREGRIDIGD  
TIVPHPTFOLSEHRSYSAKAWNRYGCVLAIELELLKDLIELDVLVYGANVQEEVGL  
RKAKASAENIDEDVAFVDCSPANDYKGSQPSLSELGKGTILIRIKDGTMLIKPVFRDY  
LKLIVBAHDIEHOYVNSPGGTGGGEIHKANIGIPTAVIGVCARYIHSTDSVFDIRDYF  
AARSLLSEALCNLDNNQIETLQYK"  
10775..11086  
/gene="MW1689"  
10775..11086  
/gene="MW1689"  
/note="ORFID:MW1689"  
/codon_start=1  
/transl_table=11  
/product="hypothetical protein"  
/protein_id="BAB95554.1"  
/db_xref="GI:21204859"  
/translation="MTKLKVIPTIIIAVAIVIIISTSIIOYINKRNPVKNVLEVK  
YFMNVKGSIVVPEFVHPEDTKVRLVYQGGITIKNGQNIHYDFADAYGEVINIVE  
R"  
complement (11211..12053)  
/gene="MW1690"  
complement (11211..12053)  
/gene="MW1690"  
/note="ORFID:MW1690"  
/codon_start=1  
/transl_table=11  
/product="conserved hypothetical protein"  
Query Match 88.1%; Score 1161; DB 1; Length 272850;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1311; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
Qy 1 ATGCACACTATCATTTTTCGGAATTAAGGTTCTGGCATGAGTTCATTAGCACAAATC 60  
Db 3785 ATGCACACTATCATTTTTCGGAATTAAGGTTCTGGCATGAGTTCATTAGCACAAATC 3726  
  
Qy 61 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATTCGATTCGATTCGATTCAGAA 120  
Db 3725 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATTCGATTCGATTCAGAA 3666  
  
Qy 121 GTTGCTCTTAGAATAGGGGATAAATAATTAACCTTGTGATTAACATTAACATAAAGAA 180  
Db 3665 GTTGCTCTTAGAATAGGGGATAAATAATTAACCTTGTGATTAACATTAACATAAAGAA 3606  
  
Qy 181 GATATGTTAGTTATACAGGTAATGCAATCCGAGTAGCCATGAAGAAATAGTACGTGCA 240  
Db 3605 GATATGTTAGTTATACAGGTAATGCAATCCGAGTAGCCATGAAGAAATAGTACGTGCA 3546  
  
Qy 241 CATCAATGAAATTAGATGTTGAAGTTATAAGTATTAATGTTTTTTAGGACAGATTTATGATCAA 300  
Db 3545 CATCAATGAAATTAGATGTTGAAGTTATAAGTATTAATGTTTTTTAGGACAGATTTATGATCAA 3486
```

RESULT 8
AF034076
LOCUS
DEFINITION

Staphylococcus aureus UDP-N-acetyl-muramoyl-L-alanine synthetase
1314 bp DNA linear BCT 16-JUL-1999

```
301 TATACCTCAGTAGCTGTAACTGGTGCAATGGTAAACCTTCTACACAGGTTTATTATCA 360  
3485 TATACCTCAGTAGCTGTAACTGGTGCAATGGTAAACCTTCTACACAGGTTTATTATCA 3426  
  
Qy 361 CATGTTATGAATGGTATAAAGACTTCATTTTAAATTTGGTGAGGACAGATATGGGA 420  
3425 CATGTTATGAATGGTATAAAGACTTCATTTTAAATTTGGTGAGGACAGATATGGGA 3366  
  
Qy 421 TTGCTCTGAAAGTGATTTATTTGCTTTTGGGCAATCGAATATAGACGTCACATTTTAAAGT 480  
3365 TTGCTCTGAAAGTGATTTATTTGCTTTTGGGCAATCGAATATAGACGTCACATTTTAAAGT 3306  
  
Qy 481 TATAAACCTGATACGCAATTAAGCAAAATATGATTTTCGATCACTCCTGATTTATTA 540  
3305 TATAAACCTGATACGCAATTAAGCAAAATATGATTTTCGATCACTCCTGATTTATTA 3246  
  
Qy 541 GATATTAATGATGTTTTTGTGATGCCATTCAGAAATGGCACATAATGTTAAABAAGGTATT 600  
3245 GATATTAATGATGTTTTTGTGATGCCATTCAGAAATGGCACATAATGTTAAABAAGGTATT 3186  
  
Qy 601 ATTGCTTGGGGTGATGATGAACATCTAGTAAATTTGAAGCAGATGTTCCAAATTTATTAT 660  
3185 ATTGCTTGGGGTGATGATGAACATCTAGTAAATTTGAAGCAGATGTTCCAAATTTATTAT 3126  
  
Qy 661 TATCGATTTAAAGATTCCGATTCACATTTATGCTCAAAATATTTCAAATACGGATAAAGT 720  
3125 TATCGATTTAAAGATTCCGATTCACATTTATGCTCAAAATATTTCAAATACGGATAAAGT 3066  
  
Qy 721 ACTGCTTTGATGTTATGTTGGATGGTGGATTTTATGATCACTTCTGCTCCACAATATT 780  
3065 ACTGCTTTGATGTTATGTTGGATGGTGGATTTTATGATCACTTCTGCTCCACAATATT 3006  
  
Qy 781 GGTGACCATACAGTTTTTAAATGCTAGTAAATTTGGATTTAGTATTTATTTAGAGAAGCTA 840  
3005 GGTGACCATACAGTTTTTAAATGCTAGTAAATTTGGATTTAGTATTTATTTAGAGAAGCTA 2946  
  
Qy 841 GATGTTACAAATATTAAGAGAGCATTAGAAAGTTGGTGGTGTGTTAAACGTCGTTTCAAT 900  
2945 GATGTTACAAATATTAAGAGAGCATTAGAAAGTTGGTGGTGTGTTAAACGTCGTTTCAAT 2886  
  
Qy 901 GAAACTACAAATGCAAAATCAAGTTTATTTAGATGATTTATGCACATCCAGAGAGAAAT 960  
2885 GAAACTACAAATGCAAAATCAAGTTTATTTAGATGATTTATGCACATCCAGAGAGAAAT 2826  
  
Qy 961 AGTGTCAAAATGAAACAGCAGAAAGAAATATCCACATAAAGAGTTCGTCAGATTTT 1020  
2825 AGTGTCAAAATGAAACAGCAGAAAGAAATATCCACATAAAGAGTTCGTCAGATTTT 2766  
  
Qy 1021 CAACACACACTTTCTCTAGAACACAGGCATTTTAAATGAAATTTGCAGAAAGTTTAAAGT 1080  
2765 CAACACACACTTTCTCTAGAACACAGGCATTTTAAATGAAATTTGCAGAAAGTTTAAAGT 2706  
  
Qy 1081 AAACAGATCGTGTATTTCTTAATGAAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1140  
2705 AAACAGATCGTGTATTTCTTAATGAAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 2646  
  
Qy 1141 TTAACGATACAGATTTAAATTTGATTAATTTCAAGTGCATCGTTAAATTAATGAAGTCT 1200  
2645 TTAACGATACAGATTTAAATTTGATTAATTTCAAGTGCATCGTTAAATTAATGAAGTCT 2586  
  
Qy 1201 ATTAATGTTATTAGAACAAATTTGATAATGCTGTTATTTTATTTATGGGTGCAGTGATATT 1260  
2585 ATTAATGTTATTAGAACAAATTTGATAATGCTGTTATTTTATTTATGGGTGCAGTGATATT 2526  
  
Qy 1261 CAAAAATTTACAAATGTCATTTTATAGATAAATTTAGGCATCAAAAAATCGCGTTTTAA 1314  
2525 CAAAAATTTACAAATGTCATTTTATAGATAAATTTAGGCATCAAAAAATCGCGTTTTAA 2472
```

(murC) gene, complete cds.
AF034076.1 GI:2642658
Staphylococcus aureus
Staphylococcus aureus
Bacteria; Firmicutes; Bacillales; Staphylococcus.
1 (bases 1 to 1314)
Lowe, A.M. and Deresiewicz, R.L.
Cloning and sequencing of Staphylococcus aureus murC, a gene
essential for cell wall biosynthesis
DNA Seq. 10 (1), 19-23. (1999)
20029270
PUBMED
10565540
2 (bases 1 to 1314)
Lowe, A.M. and Deresiewicz, R.L.
Direct Submission
Submitted (11-NOV-1997) Channing Laboratory, Brigham and Women's
Hospital and Harvard Medical School, 181 Longwood Ave, Boston, MA
02115, USA
Location/Qualifiers
1. .1314
/organism="Staphylococcus aureus"
/mol_type="genomic DNA"
/db_xref="taxon:1280"
1. .1314
/gene="murC"
1. .1314
/gene="murC"
/function="cell wall biosynthesis"
/note="MurC; UDP-N-acetylmuramate-alanine ligase"
/codon_start=1
/transl_table=11
/product="UDP-N-acetylmuramoyl-L-alanine synthetase"
/protein_id="AAB87090.1"
/db_xref="GI:2642659"
/translation="MTHYFVIGIKSGMSSLAQIMHDLGHEVQSGSDIENYVTEVALR
NKGKILPFDANNIKEDWVVIQGNFASHEEKARAHQKLDVVSNDPFLGQIIOQYT
SVAVTGAGKSTGLLSHVWNGDKTSLIGDGTGMLPESDYPFAFCEYRHFSL
YKPYATMNIIDFHPDYFKDINDVDFAFQEMAHNVKGGIILANGDEHUKRIADVPI
YYGFKSDDDIYAQNIQLTDKGTAFDVTVDGFEYDFHFLUSPQGDHVLNALAVIAISY
LEKLDNIEKALTEFGVKRFRNETTIANQVIVDDYAHHPRETSATETARKYPHK
EVAVFQPHFPTSRTOALNEFAESLSKADRVFLCEIFGSIKRENTGALATQDLIDKIEG
ASLINEDINSVLEQFDNAVVLFKGAGDIQKLNQVLDKLGKNAF"
ORIGIN
Query Match 75.5%; Score 995; DB 1; Length 1314;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1195; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGACACATATCATTTTGTCCGGAATTAAGGTTCTGGCATGAGTTTCATTAGCACAATC 60
DB 1 ATGACACATATCATTTTGTCCGGAATTAAGGTTCTGGCATGAGTTTCATTAGCACAATC 60
QY 61 ATGATGATTTAGGACATGAAGTTCAAGCATCGGATATTGAGAACTAGCTATTTACAGAA 120
DB 61 ATGATGATTTAGGACATGAAGTTCAAGCATCGGATATTGAGAACTAGCTATTTACAGAA 120
QY 121 GTTGCTTTAGAAATAAGGGGATAAATAATATACCATTTGATGCTTAATAACATAAAGAA 180
DB 121 GTTGCTTTAGAAATAAGGGGATAAATAATATACCATTTGATGCTTAATAACATAAAGAA 180
QY 181 GATATGATTTAGTATACAGGTAATGATCATTCGGAGTAGCCATGAAGAAATAGTAGTGA 240
DB 181 GATATGATTTAGTATACAGGTAATGATCATTCGGAGTAGCCATGAAGAAATAGTAGTGA 240
QY 241 CATCAATTTAGAAATAGATTTGTAGTTATATAGTTATTTAGGACAGATTTATGATCAA 300
DB 241 CATCAATTTAGAAATAGATTTGTAGTTATATGATTTTATAGGACAGATTTATGATCAA 300
QY 301 TATACCTTCAGTAGCTGAATCGTGCACATCGTGAACCTTCTACAAAGGTTTATATCA 360
DB 301 TATACCTTCAGTAGCTGAATCGTGCACATCGTGAACCTTCTACAAAGGTTTATATCA 360



QY 361 CATGTTATGAATGGTATGATAAAAAGACTTCATTTTAAATGGTGATGGCACAGGATGCGA 420
DB 361 CATGTTATGAATGGTATGATAAAAAGACTTCATTTTAAATGGTGATGGCACAGGATGCGA 420
QY 421 TTGCTCGAAGATGATTAATTTTCGCTTTTGGGATGATGATGATGATGATGATGATGATGAT 480
DB 421 TTGCTCGAAGATGATTAATTTTCGCTTTTGGGATGATGATGATGATGATGATGATGATGAT 480
QY 481 TATAAACCTGATTAACCAATTAATGACAAATATGATTTTCCGATCATCTCGATTTTAA 540
DB 481 TATAAACCTGATTAACCAATTAATGACAAATATGATTTTCCGATCATCTCGATTTTAA 540
QY 541 GATATTAATGATGATTTTTCGATGATTTTCAAGAAATGACATATGATTTTAAAAAGGTTAT 600
DB 541 GATATTAATGATGATTTTTCGATGATTTTCAAGAAATGACATATGATTTTAAAAAGGTTAT 600
QY 601 ATTGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 ATTGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 TATGGATTTAAAGATTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 TATGGATTTAAAGATTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 ACTGCTTTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 ACTGCTTTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 GGTGACCATACAGCTTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 GGTGACCATACAGCTTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 GATGTTACAAATTAAGAAAGATTTAGAAACGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGT 900
DB 841 GATGTTACAAATTAAGAAAGATTTAGAAACGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGT 900
QY 901 GAAATACAAATTCGAAATCAAGTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 GAAATACAAATTCGAAATCAAGTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 AGTGCTACAAATTTGAAACAGACAGCAAGAAATATCCATTAAGAAAGTTTGTGAGTATTT 1020
DB 961 AGTGCTACAAATTTGAAACAGACAGCAAGAAATATCCATTAAGAAAGTTTGTGAGTATTT 1020
QY 1021 CAACACACATCTTCTCTAGAACACAGGCAATTTTAAATGAAATTTGCGAAGATTTTAAAT 1080
DB 1021 CAACACACATCTTCTCTAGAACACAGGCAATTTTAAATGAAATTTGCGAAGATTTTAAAT 1080
QY 1081 AAAGCAGATCGTGTATTTCTTATGTAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1140
DB 1081 AAAGCAGATCGTGTATTTCTTATGTAATTTTGGATCAATTTAGAGAAATATCTGGCGCA 1140
QY 1141 TTAACGATCAAGATTTAAATTTGATAAAATTTGAAAGTGCATCGTTAAATTAATGAAGATTC 1199
DB 1141 TTAACGATCAAGATTTAAATTTGATAAAATTTGAAAGTGCATCGTTAAATTAATGAAGATTC 1199

RESULT 9

AX191741

LOCUS

AX191741

DEFINITION

Sequence 23 from Patent WO0149775.

ACCESSION

AX191741

VERSION

AX191741.1

KEYWORDS

Staphylococcus aureus

ORGANISM

Staphylococcus aureus

REFERENCE

1

AUTHORS

Iversen, P.L.

TITLE

Antisense antibacterial cell division composition and method

JOURNAL

Patent: WO 0149775-A 23 12-JUL-2001;

Avi Biopharma, Inc. (US)

1314 bp DNA linear PAT 15-AUG-2001

FEATURES		Location/Qualifiers	
source		1. .1314	
		/organism="Staphylococcus aureus"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:1280"	
ORIGIN			
QY	1	ATGACACACTATCACTTTTGTGGAAATAAAGGTTCTGGCATCAGTTTCATTAGCACAAATC	60
Db	1	ATGACACACTATCACTTTTGTGGAAATAAAGGTTCTGGCATCAGTTTCATTAGCACAAATC	60
QY	61	ATGCATGATTAGGACATGAAGTTCACAGATCGGATATTGGAACTAGCTATTACAGAA	120
Db	61	ATGCATGATTAGGACATGAAGTTCACAGATCGGATATTGGAACTAGCTATTACAGAA	120
QY	121	GTTCCTCTTAGAATAAGGGGATAAAATATATACCAATTTGATGCTTAATAACATAAAGAA	180
Db	121	GTTCCTCTTAGAATAAGGGGATAAAATATATACCAATTTGATGCTTAATAACATAAAGAA	180
QY	181	GATATGTTAGTTATACAAAGGTAATGCAATTCGCGATAGCCATGAAGTAATAGTACGTCA	240
Db	181	GATATGTTAGTTATACAAAGGTAATGCAATTCGCGATAGCCATGAAGTAATAGTACGTCA	240
QY	241	CATCAATTTGAAATAGAGTTGTAAGTTATATATGATTTTATGAGACAGATTATTGATCAA	300
Db	241	CATCAATTTGAAATAGAGTTGTAAGTTATATATGATTTTATGAGACAGATTATTGATCAA	300
QY	301	TATACCTCAGTAGCTGTAACCTGGTGACATGGTAAACTTCTCAACAGGTTTATTATCA	360
Db	301	TATACCTCAGTAGCTGTAACCTGGTGACATGGTAAACTTCTCAACAGGTTTATTATCA	360
QY	361	CATGTTATGAATGTTGATAAAAGACTTCAATTTTAAATTTGTTGATGTCACAGGTTATGGA	420
Db	361	CATGTTATGAATGTTGATAAAAGACTTCAATTTTAAATTTGTTGATGTCACAGGTTATGGA	420
QY	421	TTGCTGAAAGTGAATTTTCGCTTTTGGAGCATGTGAATATAGACGTCACTTTTAAAGT	480
Db	421	TTGCTGAAAGTGAATTTTCGCTTTTGGAGCATGTGAATATAGACGTCACTTTTAAAGT	480
QY	481	TATAAACCTGATTACGAATATGACAAATATGATTTGATTCGATCATCTCGATTATTAA	540
Db	481	TATAAACCTGATTACGAATATGACAAATATGATTTGATTCGATCATCTCGATTATTAA	540
QY	541	GATATTAAATGATGTTTTCGATGATTCACAAATATGATTTGATTCGATCATCTCGATTATT	600
Db	541	GATATTAAATGATGTTTTCGATGATTCACAAATATGATTTGATTCGATCATCTCGATTATT	600
QY	601	ATTGCTTGGGTCATGATGAACATCTACGTAAATTTGAAGCATGTTTCCAAATTTATTAT	660
Db	601	ATTGCTTGGGTCATGATGAACATCTACGTAAATTTGAAGCATGTTTCCAAATTTATTAT	660
QY	661	TATGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTTACGATAAAGT	720
Db	661	TATGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTTACGATAAAGT	720
QY	721	ACTGCTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	780
Db	721	ACTGCTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	780
QY	781	GATGACCATACAGTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT	840
Db	781	GATGACCATACAGTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT	840
QY	841	GATGTTACAAATTTAAAGACATTTAGAAACGTTTGGTGGTGGTGGTGGTGGTGGTGGT	900
Db	841	GATGTTACAAATTTAAAGACATTTAGAAACGTTTGGTGGTGGTGGTGGTGGTGGTGGT	900
QY	901	GAACTACAAATTTGCAATCAAGTTATTGATGATGATGATGATGATGATGATGATGATGAT	960
Db	901	GAACTACAAATTTGCAATCAAGTTATTGATGATGATGATGATGATGATGATGATGATGAT	960
Db	901	GAACTACAAATTTGCAATCAAGTTATTGATGATGATGATGATGATGATGATGATGATGAT	960
QY	961	AGTGCTACATTTGAAACAGCACGAAAGAAATATCCATATAAGAGTTGTTGCGATATTT	1020
Db	961	AGTGCTACATTTGAAACAGCACGAAAGAAATATCCATATAAGAGTTGTTGCGATATTT	1020
QY	1021	CAACACACACTTTCTCTAGAACACAGGCAATTTTAAATGAATTTGAGAAAGTTTAAAGT	1080
Db	1021	CAACACACACTTTCTCTAGAACACAGGCAATTTTAAATGAATTTGAGAAAGTTTAAAGT	1080
QY	1081	AAAGCAGATCGTATTTCTATGTAATTTTGGATCAATTTAGAGAAATTTAGTGGCA	1140
Db	1081	AAAGCAGATCGTATTTCTATGTAATTTTGGATCAATTTAGAGAAATTTAGTGGCA	1140
QY	1141	TTAACCATACAGATTTTAAATTTGAAAGTTCGATCGTTAAATTAATGAGATTC	1199
Db	1141	TTAACCATACAGATTTTAAATTTGAAAGTTCGATCGTTAAATTAATGAGATTC	1199
RESULT 10			
E35613			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
OS Unidentified			
PN JP 1999225773-A/1			
PF 24-AUG-1999			
PR 03-JUL-1998 JP 1998225115			
PR 03-JUL-1997 US 60/052720			
PI NICOLA G WALLIS, MARTIN K R BURNHAM			
PC C12N15/09, A61K31/70, A61K38/53, A61K38/55, A61K39/085, A61K39/395,			
PC A61K39/395,			
PC A61K48/00, A61K48/00, C07K16/40, C12N1/21, C12P21/02, C12P21/08, PC			
C1201/68,			
PC G01N33/566, (C12N15/09, C12R1:445), (C12N1/21, C12R1:445), PC			
C12N15/00, A61K37/60,			
PC A61K37/64, (C12N15/00, C12R1:445)			
CC Strandedness: Double;			
CC Topology: Linear;			
FH Key			
FT Source			
Location/Qualifiers			
1. .1351			
/organism="unidentified"			
/mol_type="genomic DNA"			
/db_xref="taxon:32644"			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1042; Conservative			
0; Mismatches			
4; Indels			
0; Gaps			
QY	1	ATGACACACTATCACTTTTGTGGAAATAAAGGTTCTGGCATCAGTTTCATTAGCACAAATC	60
Db	22	ATGACACACTATCACTTTTGTGGAAATAAAGGTTCTGGCATCAGTTTCATTAGCACAAATC	81
QY	61	ATGCATGATTAGGACATGAAGTTCACAGATCGGATATTGGAACTAGCTATTACAGAA	120
Db	82	ATGCATGATTAGGACATGAAGTTCACAGATCGGATATTGGAACTAGCTATTACAGAA	141
QY	121	GTTCCTCTTAGAATAAGGGGATAAAATATATACCAATTTGATGCTTAATAACATAAAGAA	180

Db 142 GTTGCTCTTAGAATAAGGGGATAAAATATATACCATTTGGTGTAAATAACATAAAAGAA 201
QY 181 GATATGCTAGTATTACAAAGTAATGCATTCGGAGTAGCCATGAAGAATAAGTAGTGCA 240
Db 202 GATATGCTAGTATTACAAAGTAATGCATTCGGAGTAGCCATGAAGAATAAGTAGTGCA 261
QY 241 CATCAATTTGAAATAGATGTTGTAAGTTATATATATATATATATATATATATATATAT 300
Db 262 CATCAATTTGAAATAGATGTTGTAAGTTATATATATATATATATATATATATATATAT 321
QY 301 TATATCTCAGTAGCTGTAACCTGCTGCAATGCTGTAACCTGCTGTAACCTGCTGTAACCT 360
Db 322 TATATCTCAGTAGCTGTAACCTGCTGCAATGCTGTAACCTGCTGTAACCTGCTGTAACCT 381
QY 361 CATGTTATGATGATGTAATAAAGACTTCATTTTAAATGCTGATGTCGACAGGTATGGGA 420
Db 382 CATGTTATGATGATGTAATAAAGACTTCATTTTAAATGCTGATGTCGACAGGTATGGGA 441
QY 421 TTGCTCTGAAAGTCAATATTTTCTGCTTTTGGAGCATGCTGAATATAGACGTCACTTTTAA 480
Db 442 TTGCTCTGAAAGTCAATATTTTCTGCTTTTGGAGCATGCTGAATATAGACGTCACTTTTAA 501
QY 481 TATAAACCTGATTCGCAATATATGCAAAATATGCAAAATATGCAAAATATGCAAAATAT 540
Db 502 TATAAACCTGATTCGCAATATATGCAAAATATGCAAAATATGCAAAATATGCAAAATAT 561
QY 541 GATATTAATGATGTTTTTGGATGATTTTCCAAAGATGCAATATGTTTAAAGAGGTATT 600
Db 562 GATATTAATGATGTTTTTGGATGATTTTCCAAAGATGCAATATGTTTAAAGAGGTATT 621
QY 601 ATTGCTTGGGGTGATGATGCAATCTACGTAATGCAATGCAATGCAATGCAATGCAAT 660
Db 622 ATTGCTTGGGGTGATGATGCAATCTACGTAATGCAATGCAATGCAATGCAATGCAAT 681
QY 661 TATGATTTAAAGATTCGATGCAATTTTATGCTCAAAATATTCAAATATGCAATGCAAT 720
Db 682 TATGATTTAAAGATTCGATGCAATTTTATGCTCAAAATATTCAAATATGCAATGCAAT 741
QY 721 ACTGCTTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 742 ACTGCTTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 801
QY 781 GGTGACCATACAGTTTTTAAATGCAATGCTGTAATGCAATGCTGTAATGCAATGCTG 840
Db 802 GGTGACCATACAGTTTTTAAATGCAATGCTGTAATGCAATGCTGTAATGCAATGCTG 861
QY 841 GATGTTACAAATATTAAGAGCATTAGAAAGCTTTGGTGGTTTAAAGCTGTTTCAAT 900
Db 862 GATGTTACAAATATTAAGAGCATTAGAAAGCTTTGGTGGTTTAAAGCTGTTTCAAT 921
QY 901 GAAACTACAAATGCAAAATCAAGTTATGCTGATGATGATGATGATGATGATGATGAT 960
Db 922 GAAACTACAAATGCAAAATCAAGTTATGCTGATGATGATGATGATGATGATGATGAT 981
QY 961 AGTGCTCAATTTGAAACAGCAAGAAAGAAATATCCACATAAAGAGTTGTTGCAATTT 1020
Db 982 AGTGCTCAATTTGAAACAGCAAGAAAGAAATATCCACATAAAGAGTTGTTGCAATTT 1041
QY 1021 CAACCCACACTTTCTCTAGAACACA 1046
Db 1042 CAACCCACACTTTCTCTAGAACACA 1067

RESULT 11
BD178030
LOCUS BD178030 1351 bp DNA linear PAT 16-APR-2003
DEFINITION MurC.
ACCESSION BD178030
VERSION BD178030.1 GI:30015294
KEYWORDS JP 2002300888-A/1.
SOURCE unidentified
ORGANISM unclassified.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 1351)
Wallis, N.G. and Burnham, M.K.R.
MurC
Patent: JP 2002300888-A 1 15-OCT-2002;
SMITHKLINE BEECHAM CORP. SMITHKLINE BEECHAM PLC
OS Unidentified
PN JP 2002300888-A/1
PD 15-OCT-2002
PF 23-DEC-2001 JP 2001391079
PR 03-JUN-1997 US 60/052720
PI NICOLA G WALLIS, MARTIN K R BURNHAM
PC C12N15/09, A61K31/7088, A61K38/43, A61K45/00, A61K48/00,
PC A61P1/00,
PC A61P7/02, A61P9/00, A61P11/00, A61P13/02, A61P13/12, A61P17/00, PC
A61P17/02,
PC A61P19/02, A61P19/08, A61P21/00, A61P25/00, A61P27/00, A61P29/00,
PC A61P31/04,
PC A61P39/02, A61P43/00, C07K14/195, C12N1/15, C12N1/19, C12N1/21, PC
C12N5/10,
PC C12P21/02, C12Q1/02, C12Q1/68, G01N33/15, G01N33/50, C12N15/00, PC
C12N5/00,
PC A61K37/48, A61K37/64
CC MurC
FH Key Location/Qualifiers
FT source 1. .1351 /organism='Unidentified'.
FT Location/Qualifiers
1. .1351
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN

Query Match 63.9%; Score 842; DB 6; Length 1351;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1042; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGACACACTATCATTTTGTGCGAATTAAGGTTCTGGCATGAGTTCAATAGCACAAATC 60
Db 22 ATGACACACTATCATTTTGTGCGAATTAAGGTTCTGGCATGAGTTCAATAGCACAAATC 81
QY 61 ATGCATGATTTAGGACATGAGTTCAAGGATCGGATATTGAGACTAGTATTACAGAA 120
Db 82 ATGCATGATTTAGGACATGAGTTCAAGGATCGGATATTGAGACTAGTATTACAGAA 141
QY 121 GTTGCTCTTAGAATAAGGGGATAAAATATATACCATTTGATGCTTAATAACATAAAAGAA 180
Db 142 GTTGCTCTTAGAATAAGGGGATAAAATATATACCATTTGATGCTTAATAACATAAAAGAA 201
QY 181 GATATGCTAGTTATACAAAGTAATGCATTCGCGAGTAGCCATGAAGAAATAGTAGTGCA 240
Db 202 GATATGCTAGTTATACAAAGTAATGCATTCGCGAGTAGCCATGAAGAAATAGTAGTGCA 261
QY 241 CATCAATTCGAATTAGATGTTGTAAGTTATATATATATATATATATATATATATAT 300
Db 262 CATCAATTCGAATTAGATGTTGTAAGTTATATATATATATATATATATATATATAT 321
QY 301 TATATCTCAGTAGCTGTAACCTGCTGCAATGCTGTAACCTGCTGTAACCTGCTGTAACCT 360
Db 322 TATATCTCAGTAGCTGTAACCTGCTGCAATGCTGTAACCTGCTGTAACCTGCTGTAACCT 381
QY 361 CATGTTATGATGATGTAATAAAGACTTCATTTTAAATGCTGATGTCGACAGGTATGGGA 420
Db 382 CATGTTATGATGATGTAATAAAGACTTCATTTTAAATGCTGATGTCGACAGGTATGGGA 441
QY 421 TTGCTCTGAAAGTCAATATTTTCTGCTTTTGGAGCATGCTGAATATAGACGTCACTTTTAA 480
Db 442 TTGCTCTGAAAGTCAATATTTTCTGCTTTTGGAGCATGCTGAATATAGACGTCACTTTTAA 501
QY 481 TATAAACCTGATTCGCAATATATGCAAAATATGCAAAATATGCAAAATATGCAAAATAT 540
Db 502 TATAAACCTGATTCGCAATATATGCAAAATATGCAAAATATGCAAAATATGCAAAATAT 561

Db 507 ATTAGACAAATTTCGATAATGCTGTATTTT 478

Search completed: June 6, 2004, 06:04:07
Job time : 5363 secs